Thesis for the degree of Doctor of Philosophy

Dynamic Structure Discovery and Ion Transport in Liquid Battery Electrolytes

Rasmus Andersson

Department of Physics
Chalmers University of Technology
Gothenburg, Sweden 2020

Dynamic Structure Discovery and Ion Transport
in Liquid Battery Electrolytes
RASMUS ANDERSSON
ISBN 978-91-7905-413-7

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie nr 4880
ISSN 0346-718X
Department of Physics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31-772 1000

Cover:
Structure detected by dynamic structure discovery, with expanded local structures. The electrolyte is a highly concentrated mixture of LiTFSI and LiTDI in G4 (paper II). Above the scene hovers the logo of CHAMPION, the software implementing the methods presented in this thesis.

This work has received funding through the HELIS project (European Union's Horizon 2020 research and innovation program under Grant Agreement No. 666221), the Swedish Energy Agency (#P43525-1 and #P39909-1), and Chalmers Area of Advance: Materials Science, Theory and Modelling scheme of Advanced User Support.

Chalmers Digitaltryck
Gothenburg, Sweden 2020

Abstract

The lithium-ion battery (LIB), the realisation of which earned the Nobel Prize in Chemistry 2019, has since its 1991 commercialisation become the dominant energy storage technology first for cell phones and other mobile electronics, then for power tools and other domestic appliances, and currently for electric cars and other vehicles. However, many applications would still benefit from higher power and energy densities, longer life-lengths and safer batteries. Such improvements would for example accelerate the electrification of transport, lower the pollution and the greenhouse gas emissions.

Electrolytes are extremely crucial for the operation of the LIBs, yet they have so far changed surprisingly little the last 25 years. Further improvement can be made by novel electrolyte concepts. Highly concentrated electrolytes (HCEs) may enable higher energy and power densities, as well as improved thermal, chemical and electrochemical stabilities as compared to the current state-of-the-art, while also being more flexible in their composition. They also have more complex structures and ion transport mechanisms. I here present a novel method for studying both more standard electrolytes and HCEs by analysing molecular dynamics simulation trajectories. This method automatically detects the time-dependent structures present and characterises them by statistical physics, giving an extraordinarily detailed view of the structure and dynamics. I describe the theory and implementation of this method as well as its application to several HCEs and the ubiquitous LP30 electrolyte. These studies enhance the picture of ion transport conveyed previously and future application should add substantially to the design of battery electrolytes and beyond.

Keywords: lithium-ion batteries, electrolytes, ion transport mechanisms, dynamic structure discovery, statistical physics

List of Papers

This thesis is based on the following papers:

I *Ion Transport Mechanisms via Time-dependent Local Structure and Dynamics in Highly Concentrated Electrolytes*
R. Andersson, F. Årén, A. A. Franco and P. Johansson
J. Electrochem. Soc. **167** 140537 (2020)
doi.org/10.1149/1945-7111/abc657

II *Designing High-Performant Lithium Battery Electrolytes by Utilizing Two Natures of Li^+ Coordination: LiTDI/LiTFSI in Tetraglyme*
P. Jankowski, **R. Andersson** and P. Johansson
Batt. & Supercaps, accepted Sep 10th 2020
doi.org/10.1002/batt.202000189

III *CHAMPION: Chalmers Hierarchical Atomic, Molecular, Polymeric & Ionic Analysis Toolkit*
R. Andersson, F. Årén, A. A. Franco and P. Johansson
Submitted

IV *Structure, Dynamics and Ion Transport Mechanisms in Lithium-Ion Battery Electrolytes: Origin and Onset of Highly Concentrated Electrolyte Behavior*
R. Andersson, F. Årén, A. A. Franco and P. Johansson
Submitted, under review

V *Structure and Dynamics of Ion Transport in the Ubiquitous Lithium-ion Battery Electrolyte LP30*
R. Andersson, O. Borodin and P. Johansson
Manuscript

The articles are reprinted with permissions from the publishers.

Other results arising from my thesis work:

- Swedish patent application #2051245-5
- The start-up Compular AB, co-founder

Contribution Report

My most substantial contribution to all included papers has been the design, development and implementation in our software CHAMPION of the methods for dynamic structure discovery and subsequent analysis.

I I designed the study in collaboration with my co-authors. I set up and performed the AIMD simulations together with FÅ. I performed the CHAMPION analysis, analysed all results and wrote the first draft of the paper. I co-authored the full paper.

II I performed the CHAMPION analysis, analysed the results and wrote the first draft for the parts of the paper based on the CHAMPION analysis. I gave feedback on the full paper to my co-authors.

III I had the original ideas of the methods presented. I designed, developed and implemented them in code in collaboration with FÅ and under supervision of PJ and AAF. I wrote the first draft of the paper and co-authored the full paper.

IV I designed the study in collaboration with my co-authors. I set up and performed the AIMD simulations together with FÅ. I performed the CHAMPION analysis, analysed all results, wrote the first draft of the paper and co-authored the full paper.

V I designed the study in collaboration with my co-authors. I performed the CHAMPION analysis, analysed all results and wrote the first draft of the Results and Discussion section. I co-authored the full paper.

List of Acronyms

ACN	acetonitrile
AIMD	*ab initio* MD
AM	Active Material
CHAMPION	Chalmers Hierarchical, Atomic, Molecular, Polymeric & Ionic Analysis Toolkit
CLT	Central Limit Theorem
CN	Coordination Number
CPMD	Car-Parrinello MD
DFT	Density Functional Theory
DMC	dimethyl carbonate, $C_3H_6O_3$
DME	dimethoxy ethane $C_4H_{10}O_2$
DOL	1,3-dioxolane, $(CH_2)_2O_2CH_2$
DSD	Dynamic Structure Discovery
EC	ethylene carbonate, $C_3H_4O_3$
EMC	ethyl methyl carbonate, $C_4H_{10}O_3$
ESW	Electrochemical Stability Window
FSI	bis(fluorosulfonyl)imide, $[N(SO_2F)_2]^-$
G4	tetraglyme, $C_{10}H_{22}O_5$
GA	Genetic Algorithm
HCE	Highly Concentrated Electrolyte
LCO	$LiCoO_2$
LFP	$LiFePO_4$
LIB	Lithium-Ion Battery
LP30	1 M $LiPF_6$ in EC/DMC, 1:1 by volume
MD	Molecular Dynamics
ML	Machine Learning
MSM	Multi-Scale Modelling
NCA	$LiNi_xCo_yAl_zO_2$, where $x + y + z = 1$
NMC	$LiNi_xMn_yCo_zO_2$, where $x + y + z = 1$
PBC	Periodic Boundary Conditions
PBE	Perdew-Burke-Ernzerhof
PC	propylene carbonate, $C_4H_6O_3$
PS	polysulfide
pCN	partial CN
pSN	partial SN
SEI	Solid-Electrolyte Interphase
SN	Solvation number
SIL	Solvated Ionic Liquid
SoC	State-of-Charge
TDI	2-trifluoromethyl-4,5-dicyanoimidazole, $[C_6F_3N_4]^-$
TFSI	bis(trifluoromethanesulfonyl)imide, $[N(SO_2CF_3)_2]^-$

Table of Contents

Chapter 1

Introduction

High energy density rechargeable batteries can contribute to the transition to a CO_2 neutral economy, not the least for hybrid and fully electric vehicles. Electrifying transport is crucial for climate change mitigation since the transport sector stands for about 15% of global CO_2 emissions [1].

Since their commercialization in the early 1990's, lithium-ion batteries (LIBs) have revolutionised portable electronics, and their volumetric capacity has more than tripled, from 200 Wh/l for the first commercial LIB to 650 Wh/l for energy optimised cells today [2]. Due to high energy and power densities, as well as long cycle and calendar lives, LIBs have become the most prevalent energy storage technology for fully electric vehicles [3], and in 2019 key discoveries behind the development of LIBs were rewarded the Nobel Prize in Chemistry [4].

The electrolytes of today's state-of-the-art LIBs, however, are volatile and flammable, which makes LIBs prone to thermal runaway under abuse or extraordinary conditions [5]. These safety issues are compounded by the large amounts of energy contained in electric vehicle battery packs compared to handheld electronics. The electrolytes are also not electrochemically stable towards novel high voltage positive electrodes [6], while highly optimised in their formulation and not substantially variable without losing part of their functionality [5]. Creating LIB cells that combine both higher energy density with improved safety thus require more or less entirely new electrolyte designs.

1.1 Scope

This thesis aims to contribute to the understanding of complex liquid battery electrolytes by means of a novel modelling framework, and its implementation CHAMPION: Chalmers Hierarchical, Atomic, Molecular, Polymeric & Ionic Analysis Toolkit, based on dynamic structure discovery (DSD) in simulated atomic trajectories of bulk electrolyte, and subsequent classification and

analysis of the discovered structures.

The analysis covers foremost highly concentrated electrolytes (HCEs) and liquid LIB electrolytes - as also reported in papers I-II and IV-V, but also the prospects to model lithium-sulfur (Li-S) battery electrolytes and to use the method as a layer in multi-scale modelling (MSM) are briefly outlined - as it was an initial path of research within the thesis work.

1.2 Lithium-Ion Batteries

An LIB cell consists mainly of three parts: a positive electrode, a negative electrode and a separator wet by a liquid electrolyte (Fig. 1.1). The electrodes are sometimes referred to as anode (negative) and cathode (positive), but this designation is strictly true only when discharging the cell, why henceforth they will be referred to as positive and negative electrodes [7].

Both electrodes consist of particles of ion and electron conducting active materials (AMs) creating a porous structure held together by a binder material, enhancing the mechanical properties and most often an additive enhancing the electron conduction – often carbon-based. Each electrode is cast on a current collector foil that connects the electrode AM particles to each other and to the external circuit and give additional mechanical stability. For the positive electrode Al is used, while for the negative electrode, heavier and more expensive Cu must be used as Li alloys with Al at low electrochemical potentials [8].

Both the AMs in LIBs are intercalation compounds, *i.e.*, they allow Li^+ ions to be inserted at specific sites in the electrode matrix without significantly changing the structure of the host material – intercalated. The negative electrode AM is usually graphite, which can host Li^+ ions between its graphene layers, while the positive electrode AMs are more diverse, but in general transition metal oxides or phosphates, *e.g.* $LiCoO_2$ (LCO), $LiFePO_4$ (LFP), $LiNi_xMn_yCo_zO_2$(NMC), or $LiNi_xCo_yAl_zO_2$ (NCA), where in the two last cases $x + y + z = 1$. All these have either layers or channels allowing Li^+ intercalation and transport [8].

The two electrodes are sandwiched around a separator, which is a microporous, electronically insulating, often polymeric material, whose function is to avoid electronic short circuit of the cell. The separator is wet by a liquid electrolyte, which conducts ions in order to balance the electron transport in the external circuit. [8]

In the discharged state, the positive electrode is lithiated and the negative electrode delithiated. Upon charging, the applied voltage oxidizes the

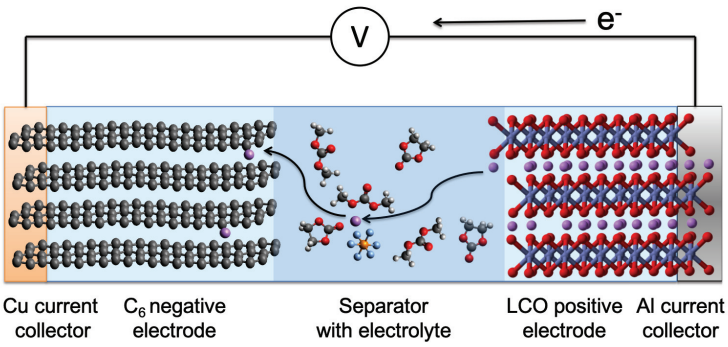


Figure 1.1: Schematic of a LIB in the beginning of the charging process.

transition metal atoms in the positive electrode and releases electrons to the external circuit. The change in oxidation state leads to release of intercalated Li^+ ions into the electrolyte, while the electrons migrate along the external circuit to the negative electrode which is reduced and accepts Li^+ ions (Fig. 1.1).

The energy content of an electrochemical cell is determined by the voltage between the electrodes, $V = E^+ - E^-$ where E^+ and E^- are the electrochemical potentials of the positive and the negative electrodes, respectively, *vs.* some reference potential, and the amount of charge transferred between them during a full discharge. The total energy is given by

$$E = \int_0^C V(Q)dQ, \tag{1.1}$$

where $V(Q)$ is the voltage after Q amount of charge has passed from the negative to the positive electrode and the integral runs from 0 to the full capacity of the cell, C, defined as the total amount of charge that can be reversibly transferred [7]. The (volumetric) energy density of the cell is given by the total energy per volume and the specific energy, or gravimetric energy density, by the energy per mass, where the mass or energy can be calculated on the AM, the cell or the battery level.

An important property of an AM is thus its specific capacity, *i.e.*, its capacity per unit mass. The theoretical specific capacity is given by an amount of charge divided by the mass of AM required to hold that charge. Table 1.1 shows average potentials and theoretical specific capacities for a number of common electrode materials in LIBs and Li-S batteries.

Table 1.1: Average potentials and theoretical specific capacities for some common AMs in LIBs [9] and Li-S batteries [10]

Active Material	Average Potential *vs.* Li^+/Li^o (V)	Specific Capacity (mAh/g)
Li	0.0	3860
Graphite	0.1	372
S	2.3	1670
LFP	3.4	170
$LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$	3.7	278
LCO	4.0	274

1.3 Lithium-Sulfur (Li-S) Batteries

Li-S batteries are one of the more researched next generation battery (NGB) concepts, in its simplest design based on a S positive and a Li metal negative electrode [11, 12]. The Li-S battery concept is attractive due to the (very) low cost of S as an AM and the very high theoretical capacities of both AMs compared to those used in LIBs (Table 1.1). However, the voltages of Li-S cells tend to be considerably lower than for LIBs (Table 1.1, [11]).

Sulfur exists in several allotropes, the most common is S_8 rings [13]. Sulfur can in a Li-S cell be reduced in a number of steps, resulting in various lithium-polysulfides (PSs) [14]:

$$S_8 \rightarrow Li_2S_6^{2-} \rightarrow \ldots \rightarrow Li_2S. \tag{1.2}$$

The most obvious problem with using sulfur as an electrode is its negligible electronic conductivity. Therefore carbon/sulfur (C/S) composites with relatively high concentrations of carbon, up to 50 wt%, are used, but this then lowers the practical cell energy density achievable significantly [11].

During discharge electrons travel from the Li metal, which is oxidized and therefore stripped of Li, to the C/S composite, where S_8 rings are reduced and react to form various Li-PSs. At decreasing state-of-charge (SoC), the average PS chain length decreases until short-chain PSs are finally reduced to Li_2S. During charge, the reverse process happens.

One of the main problems of realising practical Li-S batteries follows from the solubility of elemental sulfur and the intermediate long and medium chain length PSs into the electrolyte [11, 14]. When S and PSs are transported away from the electrode, AM is lost and they might eventually end up at the

surface of the negative electrode and partake in parasitic reactions, all leading to irreversible capacity losses. One way to counter this PS shuttle mechanism is to engineer electrolytes to have low PS solubility, while maintaining good ionic transport [14]. Rational design of Li-S electrolytes to this end requires a very good understanding of the details of the transport mechanisms.

1.4 Electrolytes

Typical electrolytes for Li batteries, both LIBs and Li-S batteries, consist of a lithium salt dissolved in a polar solvent [5]. Recent guidelines by Flamme et al. lists the following desiderata for electrolytes in high energy density LIBs [15]:

- ionic conductivity above a few mS/cm,
- electrochemical stability window (ESW) $\geq$ 4.5 V *vs.* Li^+/Li^o,
- liquidus range at least between -20°C to 180°C,
- relative permittivity $\geq$ 20,
- chemically inert against all cell components to avoid unwanted side reactions,
- good wettability towards electrodes and separator.

Electrolytes should also preferably have high flash point, low toxicity, low cost and be environmentally benign.

1.4.1 Conventional LIB Electrolytes

The large potential gap between the electrodes in LIBs prohibits aqueous electrolytes from being used, as water has a quite narrow ESW, 1.23 V [16], outside which hydrogen evolution due to reduction or oxygen evolution due to oxidation takes place. Instead, mixtures of aprotic, organic solvents, typically cyclic and linear alkyl carbonates are used, due mainly to their high oxidative stability, *ca.* 4.5-5.0 V *vs.* Li^+/Li^o [17], where the cyclic carbonates provide high relative permittivities needed for high salt solubility while the linear carbonates lower the viscosity, thus increasing ionic conductivity [5].

In fact, the commonly used electrolytes, while having significantly wider ESWs than aqueous electrolytes before breaking down due to oxidation or

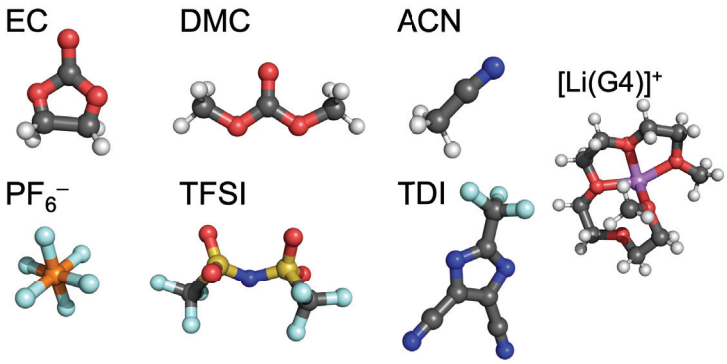


Figure 1.2: Solvents and anions studied in this thesis.

reduction, are not thermodynamically stable towards reduction against commonly used LIB negative electrodes. Modern LIBs therefore depend on reduction of solvent molecules and anions during the first cycle(s) in order to form a layer, the solid electrolyte interphase (SEI), between the electrode and the electrolyte, protecting the latter from continual decomposition, while allowing diffusion of Li ions through [18, 19].

The most used electrolyte formulation for LIBs, LP30, consists of 1 M $LiPF_6$ dissolved in an equal volume mixture of the cyclic ethyl carbonate (EC) and the linear dimethyl carbonate (DMC) [5, 20–23] (Fig. 1.2). The 1 M salt concentration is chosen because it gives the maximum ionic conductivity; lower salt concentration leads to fewer charge carriers, while higher salt concentration leads to higher cation-anion aggregation neutralizing a fraction of the charge carriers and furthermore also increases the viscosity. DMC can be replaced by similar carbonates, such as ethyl methyl carbonate (EMC), and the relative concentrations of linear and cyclic carbonates can be varied to make different trade-offs between properties, but so far both EC and PF_6^- have been found hard to replace. EC forms stable SEIs without co-intercalating into graphite and causing exfoliation of graphitic sheets (as propylene carbonate (PC) does) [24] while PF_6^- has a good balance of properties not matched by any other known anion, although it is thermally unstable and very reactive towards traces of water [5]. Most importantly its decomposition products stabilize the Al current collector by creating a thin AlO_xF_y layer [5]. For the standard electrolyte the main available degree of freedom in composition is the addition of small (typically $\leq$ 5 wt%) concentrations of functional additives, *e.g.* flame retardants, overcharge protectors and SEI-formers [25–30].

1.4.2 Highly Concentrated Electrolytes (HCEs)

HCEs, first addressed in the battery literature already in 1985 [31], has attracted much renewed interest in recent years [32–57].

HCEs are usually similar in composition to conventional elctrolytes with the difference that the salt concentration is comparable to the concentration of solvent in terms of molar ratio, resulting in *ca.* 3-5 M electrolytes. This much higher concentration modifies the local structure by making the solvent too scarce to fully solvate the cations, inviting a greater degree of ion-ion interaction. It also affects the electrochemical stability, as the local structure determines the vulnerability to oxidation/reduction, and thus determines both the size of the ESW as well as the electrolyte decomposition products [35], which in turn affects the ability to passivate the electrodes. Increasing the electrolyte salt concentration has been demonstrated to enable reversible cycling at both lower [31, 35, 37] and higher [38, 39, 46] potentials than the corresponding less concentrated electrolytes. HCEs can also be varied almost endlessly in composition.

A special class of HCEs arising out of Watanabe's group, solvate ionic liquids (SILs) consist of equimolar mixtures of Li^+ salts and chelating solvents such as glymes, *e.g.* tetraglyme (G4) where the Li^+ ions are at all times wrapped inside a solvent molecule, effectively making it a larger, softer cation $[Li(G4)]^+$ [39, 58, 59] (Fig. 1.2). The complete lack of uncoordinated solvent molecules in SILs makes this concept quite promising for expanding the ESW of the solvent, while also enabling fast Li^+ transport, as a glyme solvated ion is typically quite a lot smaller than one solvated by other common battery solvents. In paper II, we investigate electrolytes based on equimolar mixtures of salt and tetraglyme, one of which forms a SIL.

Recently, a lot of the interest in the area of HCEs have gone into two new classes of electrolytes: aqueous HCEs [16, 60–62] and localised HCEs (LHCEs) [63, 64]. In the former, extremely high salt concentrations enable expanding the ESW of aqueous electrolytes to the point that they could be part of practical, safe and potentially more environmentally friendly aqueous batteries. LHCEs start from HCEs and dilute them with a co-solvent with poor ability to dissolve the salt, in order to combine the stability gains from HCEs with the faster transport of the co-solvent. While these two electrolyte concepts are not further treated here, they are both good candidates for future application of the methods developed in this thesis.

1.4.3 Electrolytes for Li-S Batteries

Sulfur as an electrode has a much lower potential (2.3 V *vs.* Li^+/Li^o as compared to the AMs used in LIBs, and hence the selection of electrolytes for Li-S batteries is not as constrained w.r.t. oxidative stability [14]. On the other hand, the shuttling of PSs from the C/S composite to the Li metal electrode and subsequent precipitation of unsoluble PSs leads to irreversible AM loss unless the Li metal surface is passivated [10]. Therefore, the electrolytes for Li-S batteries should be designed to limit PS solubility [11, 14].

The most common electrolyte is a 1 M LiTFSI in an equimolar (1:1) solution of 1,3-dioxolane (DOL) and 1,2-dimethoxyethane (DME), with $LiNO_3$ as an additive or co-solvent (*ca.* 0.5 M) to passivate the Li surface [14] (Fig. 1.2). Both the LiTFSI and LiFSI salts can be used in Li-S batteries without any risk of Al corrosion due to the relatively low potential of sulfur.

The main qualitative difference to LIB electrolytes is that Li-S battery electrolytes contain dissolved PS species which vary in type and concentration, as functions of cycling and SoC. Hence also the physico-chemical properties of the electrolytes vary [14]. Furthermore, the extent to which the PSs are coordinated to Li ions and the extent to which they partake in Li^+ ion ligand exchange, *i.e.* both the structure and the dynamics, are complex, poorly understood phenomena [65].

1.5 Ion Transport in Liquid Electrolytes

Since the primary function of battery electrolytes is to facilitate ion transport, specifically Li^+ ion transport in the context of LIBs and Li-S batteries, the mechanisms of this transport are crucial to understand both from a scientific and technological point-of-view.

The electric field between the electrodes, due to the difference in electrochemical potential between the AMs during discharge, and due to the applied voltage during charge, is the driving force that causes net transport of ions across the separator. However, in order to understand the nature of the ion transport process it is important to appreciate the different energy scales involved. Assuming a cell voltage $V = 5$ V, the energy given to a Li^+ ion accelerated undisturbed over this voltage is 5 eV. The thermal energy at room temperature is $k_BT \approx 2.5$ keV. Thus the electrostatic bias in the direction of the field is only a negligible correction to the internal energy of the ions and molecules in the electrolyte. Given that Li^+ ions tend to stay coordinated over the ps-μs range, whereas bond vibrations occur on the fs scale, we can

furthermore conclude that the energy scale of solvation and ion interaction is additional orders of magnitude greater than the thermal energy.

Thus, while the applied voltage provides the only source of anisotropy in a liquid electrolyte, the microscopic mechanisms of transport in bulk electrolyte can be understood independent of whether or not an electric field is applied. In practice the voltage profile is very steep near the electrode surfaces why the applied voltage is more than likely important very close to the interfaces, which, however, is beyond the scope of this thesis.

The property of greatest interest to the transport performance of an LIB or Li-S electrolyte is the Li^+ ion conductivity, σ, defined by the relation

$$\boldsymbol{J}^+ = \sigma^+ \boldsymbol{E}, \tag{1.3}$$

relating the current density $\boldsymbol{J}$ of Li^+ ions resulting from an applied electric field $\boldsymbol{E}$. σ^+ is generally a tensor, but is reduced to a scalar in isotropic media such as liquid electrolytes. The Li^+ ion conductivity can be related to the diffusivity D^+ by means of the Nernst-Einstein equation, which for monovalent electrolytes (as is always the case in this work) can be written

$$\sigma^+ = \frac{cF^2}{RT} D^+ \tag{1.4}$$

(and analogously for the anion), where c is the salt concentration, F is Faraday's constant, R is the gas constant and T is the temperature. The analysis can therefore be focused on diffusivities, which is more straight-forward to compute *e.g.* from trajectories (see Section 2.4.2), than conductivities.

The simplest model for predicting the diffusivities of a species (*e.g.* a Li^+ 1^{st} solvation shell) is by the Stokes-Einstein equation [66],

$$D_i = \frac{k_B T}{6\pi\eta r_i}, \tag{1.5}$$

where η is the dynamic viscosity of the electrolyte and r_i is the hydrodynamic radius of species i.

The simplicity of Eq. (1.5) makes it a useful first approximation or null hypothesis for ion transport, which can be used in comparison with actual electrolyte behaviour. However, it assumes that all species "feel" the same friction, which is not the case in general. The friction between each pair of species depends on the nature and strength of their interaction. An improvement upon Eq. (1.5) is obtained by including friction coefficicents between each pair of species, as in Newman concentrated electrolyte theory [67–69],

which is based on the Maxwell-Stefan model for multi-component diffusion [70],

$$c_i \nabla \mu_i = \sum_j K_{ij} \left(\boldsymbol{v}_j - \boldsymbol{v}_i \right), \tag{1.6}$$

where the left-hand-side expresses the driving force for diffusion in terms of the gradient of the electrochemical potential μ_i for species i and the right-hand-side gives the resulting relative average drift velocities of each pair of species as a consequence of the friction coefficients K_{ij} between them. Combining equation (1.6) with the continuity equation gives a system of equations that can be solved for the flux of each species in the laboratory frame-of-reference.

Experimentally, the total ionic conductivity $\sigma = \sigma^+ + \sigma^-$, given by

$$\boldsymbol{J} = \left(\sigma^+ + \sigma^- \right) \boldsymbol{E}, \tag{1.7}$$

where $\boldsymbol{J}$ is the total ionic current density, is easier to measure experimentally than the Li^+ ion conductivity, and hence more often reported, in spite of its lesser importance for battery performance. Nonetheless, several studies based on Maxwell-Stefan theory, have been published that decompose the total conductivity into cationic and anionic contributions [53, 71–73].

To supplement this measure, there is a need to introduce an additional number to quantify the degree to which ion transport is cationic or anionic. Two related but distinct concepts are often used and confused in the literature: the *transport number* and the *transference number*. The cationic *transport number* is defined as the fraction of the total current that is carried by Li^+ ions under the assumption of no ion aggregation. The *transference* number is instead defined as the fraction of the migration current, *i.e.* only the current driven by an electric field, excluding any currents due to concentration gradients, that is carried by Li^+ ions, regardless of the speciation. Transference number is thus well-defined even in the case of extensive ion aggregation. Unfortunately, neither of these concepts exactly match what can be practically computed from MD simulations without a net electric field, where extensive and dynamic speciation is present. Instead we define t^+, which we rather arbitrarily denote the *transport number*, to be

$$t^+ = \frac{D^+}{D^+ + D^-}, \tag{1.8}$$

equivalent to the definition of ideal transference number in [34].The confusion and ambiguity in the definitions, measurements and models of transport and transference numbers means that comparisons between values obtained by different means are quite problematic. However, concentration trends, *etc*, seem to be quite consistent across methods.

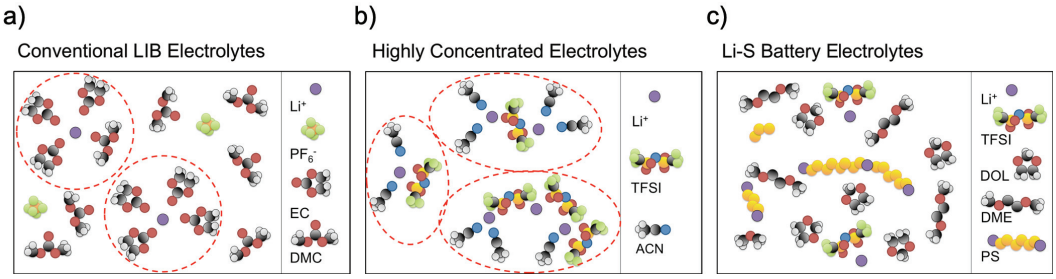


Figure 1.3: Schematic illustration of the local structures in a) conventional, b) highly concentrated and c) Li-S battery electrolytes.

1.5.1 Challenges in Modelling Ion Transport

The above outlined approach works well if the Li^+ ions are predominantly transported in long-lived solvation shells of a well-defined size – vehicular transport – but lacks explanatory power for microscopic transport mechanisms if ion hopping or solvent exchange plays important roles, which both experimental [72] and computational [74] studies suggest.

Vehicular transport leads to more mobile anions than Li^+ ions as the latter drag along their first solvation shell while a substantial fraction of the anions move independently (Fig. 1.3a). Although considerable ion-pairing seems to be the norm [41, 72], many electrolytes with 1 M salt concentration (*ca.* 1:20 salt:solvent) have been found to have low variance in coordination numbers compared to HCEs [49], and thus the former are more amenable to Newman style modelling than the latter.

With much fewer solvent molecules available per Li^+ ion than in a conventional 1 M electrolyte, the first solvation shell of the cation will be different in a HCE, no longer ruled mainly by the cation-solvent interaction. They will hence be both more complex and more varied, especially as the dearth of solvent molecules leads to a less effective Coulombic screening and in turn, to more extensive ion-ion coordination [41, 49, 72], and contain a variable number of solvent molecules and anions. The latter can also be further coordinated to other cations, creating clusters and/or networks, and hence the possible number of local structures increases drastically (Fig. 1.3b). At the same time, the ligand exchange rate changes with concentration, why different transport mechanisms may become important [42]. This may explain why t^+ increases for HCEs, but there is yet no proper understanding of the mechanism. In contrast to the vehicular transport mechanism often assumed for conventional 1 M electrolytes, a Grotthuss-like hopping mechanism can here contribute, in

which individual Li^+ ions hop from cluster to cluster [42], akin to the hopping of excess protons between water molecules through hydrogen bonds in aqueous electrolytes. However, many more complex mechanisms can easily be imagined. In Li-S batttery electrolytes dissolved S and PSs can both coordinate to Li^+ ions and affect the overall viscosity, and thereby further increase the complexity of the ion transport – both structurally and dynamically (Fig. 1.3c).

To be capable of explaining ion transport in highly concentrated and Li-S battery electrolytes any method applied must take the stance of avoiding unwarranted assumptions, why it almost must be based on modelling on the molecular level. Second, because of the structural complexity, it must be based on large enough systems to enable statistically reliable conclusions to be made regarding the ion transport mechanism(s). Third, any system modeled needs to be simulated for a long enough time (or equivalent) so that the dynamics are adequately sampled. Altogether, this combination of requirements is hard to fulfill with a single simulation method, which motivates multi-scale methodology to be applied. *Ab initio* molecular dynamics (AIMD) based on density functional theory (DFT) gives high accuracy but poor scaling with system size [75]. Classical MD on the other hand, is highly sensitive to the quality of the parameterization used for the interactions (*i.e.* the force field) [76, 77]. Unfortunately, well-known transferable force fields for organic chemistry-based systems have been found unsatisfactory in reproducing experimental transport properties in systems with high degrees of ionic association [78]. Recent works seem to indicate that a suitable scaling of the atomic charges can provide a good fit with transport properties. However, there is no guarantee that such ad-hoc measures will reproduce accurate molecular-scale dynamics.

The approach initially aimed at was to combine the strengths of both approaches through a multi-scale modelling (MSM) approach, where the force field is optimised by machine learning based on the forces observed in the AIMD simulation. This approach has, however, not yet been brought to fruition. In this thesis I both describe some of our earlier unsuccessful attempts and outline future research directions towards this end, but also how AIMD and classical MD simulations have been successfully used, depending on the systems and research questions addressed in each paper.

In addition, after having run the simulations, the complexity of these electrolytes also demand much from post-processing and analysis methods in order to qualitatively address the question of transport mechanisms. To that end, the second, and hitherto more successful leg on which this thesis stands, is a method for dynamic structure discovery (DSD). Based on atomic trajecto-

ries, this algorithm finds time-dependent bonds between atoms based on what pairs of atoms move together over time. By analysing and partitioning the global bond graph of the system using statistical physics, we can describe the structure and dynamics of a simulated system in detail, including quantifying the contributions to ion transport from different transport mechanisms, which themselves are automatically discovered by the methods.

A short mention should be given also to more advanced thermodynamic models in the tradition of the Newman approach [68, 69, 79–84]. These models can be very powerful for increasing the understanding of electrochemical systems by abstracting their essential features. This approach is almost the reverse of the one taken here. While they creatively build models based on hypotheses for the molecular-scale dynamics and test these hypotheses by whether the models reproduce the qualitative behaviour of real systems, we instead start by simulating the systems in detail and trying to make sense of the resulting data. We consider these two approaches highly complementary and two essential parts of building MSMs that are both quantitatively predictive and explanatory. They are, however, due to this same complementarity, outside the scope of this thesis and will not be discussed further.

Following this introduction, in chapter I describe the computational methods used to simulate electrolytes, and tools used in the analysis of the simulated trajectories. Thereafter, in Chapter 3, the novel methods developed within this thesis work are presented. Chapter 4 provides a holistic overview of our analysis results across different classes of electrolytes. The closing chapter gives our conclusions and points out directions for future work.

Chapter 2

Computational Background

This chapter begins by giving an overview of the simulation techniques used in this thesis, starting from the most microscopic and building up towards AIMD and classical MD simulations. Thereafter, different computational tools not directly related to the simulations are introduced together with examples of their usage in this thesis, in turn: graph theory, statistical physics and optimisation methods.

2.1 Density Functional Theory

Applying computational quantum chemistry to materials is usually reduced to predicting the ground state electronic structure of the material for a given configuration of the atomic nuclei by applying the Born-Oppenheimer approximation [85]. This allows decoupling the dynamics of the electrons from the dynamics of the nuclei, based on their vastly different time-scales due to their disparity in mass. Furthermore, it usually suffices to consider the nuclei, and often also tightly bound electrons, as classical point particles. As only the ground state electron structure is sought, it suffices to solve the time-independent Schrödinger equation (TISE). The remaining many-body problem is still very challenging, however, as this TISE is a non-separable equation for an n-dimensional wave function [86]:

$$\left[-\sum_{i}^{n}\frac{\hbar^2}{2m_e}\nabla^2-\sum_{i}^{n}eV_{\text{ions}}(\boldsymbol{r}_i)+\frac{1}{4\pi\epsilon_0}\sum_{(i,j)}^{n}\frac{e^2}{|\boldsymbol{r}_j-\boldsymbol{r}_i|}\right]\psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n)$$
$$=E\psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n). \quad (2.1)$$

Hohenberg and Kohn proved in 1964 that all ground state properties of a many-electron system are uniquely determined by the electron density [87],

$$n(\boldsymbol{r}) = \sum_{i=1}^{N} \int_{\boldsymbol{R}^3} \delta(\boldsymbol{r}_i - \boldsymbol{r}) \psi^*(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) d\boldsymbol{r}_1 d \ldots \boldsymbol{r}_n. \tag{2.2}$$

This means that there exists a functional $E[n]$ which determines the total ground state energy of the system as a function of the density. They further showed that the correct ground state electron density minimises the total energy under the constraint of conserving the number of electrons [87],

$$\int_{\boldsymbol{R}^3} n(\boldsymbol{r}) d\boldsymbol{r} = N. \tag{2.3}$$

This is a very desirable proposition since the electron density is a function only of three spatial coordinates rather than the full $3N$-dimensional configuration space, as in Eq. (2.1). This can be expressed mathematically as

$$E_0 = \int_{\boldsymbol{R}^3} eV_{\text{ions}}(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} + \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int_{\boldsymbol{R}^3} \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' d\boldsymbol{r} + G[n], \tag{2.4}$$

where the first term is the energy from the potential of the static ions, the second term is the electron-electron repulsion and $G[n]$ is a universal functional of the electron density that captures the remaining energy contributions. However, the Hohenberg-Kohn theorems only show that such a functional exists and say nothing about how to find it.

Density functional theory (DFT) became practical due to later work by Kohn and Sham in 1965 [88]. Their method starts by splitting $G[n]$ into two terms, $G[n] = T[n] + E_{\text{xc}}[n]$, where $T[n]$ is the kinetic energy functional for a system of non-interacting electrons and $E_{xc}[n]$ collects everything not captured in previous terms. This latter functional is known as the *exchange and correlation functional* and is in general not known exactly for practically interesting systems. In practice there exists a hierarchy of methods for how to approximate this functional.

The second, and most crucial, innovation by Kohn and Sham was to consider a fictional system of non-interacting particles, governed by the Schrödinger-like equation:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_{\text{eff}}(\boldsymbol{r}) \right] \phi_i(\boldsymbol{r}) = \epsilon_i \phi_i(\boldsymbol{r}), \tag{2.5}$$

where $V_{\text{eff}} = V_{\text{ions}} + V_H + V_{\text{xc}}$ includes the ionic potential, V_{ions}, the Hartree potential, V_H, expressing electron-electron interaction in a mean field sense and the remaining potential, V_{xc} due to Pauli exclusion and Coulombic interaction beyond the mean-field approximation. Eq. (2.5) is written in terms of fictitious non-interacting spin-orbitals which have no simple physical interpretation but that reproduce the correct electron density for the corresponding interacting system, so that

$$n(\boldsymbol{r}) = \sum_i \int_{\boldsymbol{R}^3} \phi_i(\boldsymbol{r})^* \phi_i(\boldsymbol{r}) d^3\boldsymbol{r}. \tag{2.6}$$

Given an electron density, the effective potential V_{eff} can be evaluated as it is a functional of the density, whereafter Eq. (2.5) can be evaluated to give the orbitals ϕ_i and the eigenvalues ϵ_i. Given the full set of orbitals an updated density can be computed using Eq. (2.6), and the cycle is repeated until self-consistency among the equations is reached, *i.e.* when the density, potentials and orbitals have converged to stable values.

Exchange and correlation functionals in DFT exist on a "Jakob's ladder" from simple and cheap to more accurate and computationally costly [89]. The need to choose a suitable level of approximation stems from the non-locality of the interactions contained in this term, where the potential at one point depends in principle on all other points in space. In practice, however, the dependence is mostly restricted to a neighbourhood of the point, which motivates the general approach of expanding the potential in terms of the local electron density and low order spatial derivatives evaluated at the point in question:

$$V_{\text{xc}}[n(\boldsymbol{r})] = V_{\text{xc}}[n(\boldsymbol{r}), \nabla n(\boldsymbol{r}), \nabla^2 n(\boldsymbol{r}), \dots] \tag{2.7}$$

The simplest type of functionals are based on the *local density approximation* (LDA), which only depend on the density at the selected point and works well when the electron density is close to homogeneous.

The next rung on the ladder also includes the gradient, and functionals on this level are known as generalised gradient approximations (GGA). Even higher order derivatives are considered in meta-GGA functionals.

The AIMD simulations done in this thesis use the exchange-correlation functional of Perdew, Burke and Ernzerhof (PBE), which is a GGA functional [90]. PBE is extensively used, especially in AIMD simulations. While likely not the most accurate functional for any system in particular, it is non-empirical, which means that it is based on physical reasoning rather than data fitting, and thus has lower risk of being biased in its predictions for dif-

ferent systems. It has demonstrated reasonable accuracy across a wide range of systems and remains very popular 25 years after its inception.

2.2 Molecular Dynamics

Molecular dynamics (MD) is a family of methods for simulating systems of atoms, molecules and/or macromolecules over time [85]. With few exceptions, the atoms are treated as classical bodies, whose dynamics are determined by the forces acting on them, based on Newtonian physics:

$$\boldsymbol{a} = \frac{\boldsymbol{F}}{m}. \tag{2.8}$$

Based on initial conditions for the position and velocity of each atom and a fixed finite timestep length Δt, Eq. (2.8) is used to propagate the system forwards in time by finite differences. There are many schemes available for this, with different accuracies, stabilities and computational costs, but a common and practical choice is the velocity Verlet scheme, where velocities are first evaluated in the middle of each timestep for updating the positions, and then updated themselves based on the forces evaluated after updating the positions [66]:

$$\boldsymbol{v}(t + \Delta t/2) = \boldsymbol{v}(t) + \boldsymbol{a}(t)\Delta t \tag{2.9}$$

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t + \Delta t/2)\Delta t \tag{2.10}$$

$$\boldsymbol{v}(t + \Delta t) = \boldsymbol{v}(t + \Delta t/2) + \frac{1}{2}\boldsymbol{a}(t + \Delta t)\Delta t \tag{2.11}$$

To simulate condensed matter systems in the bulk phase without needing to account for surface effects, a ubiquitous trick is to use periodic boundary conditions (PBCs), as is always done in this work. A cubic simulation box containing as many atoms of the system as can be simulated at reasonable cost and over sufficiently long time scales to observe the phenomena of interest is constructed. If an atom exits through one of the boundaries, it effectively enters the box again from the opposite boundary. While this introduces artificial crystallinity in the case of amorphous systems such as liquids, if the simulation box is large enough such artifacts can become negligibly small. Systematic size studies need to be performed to assess whether this is the case, which has not been done in this thesis, because performing such studies properly would require the same order of magnitude of computational resources as used in the simulations presented here. Instead the results are interpreted with some caution and allowance for finite size effects to potentially be important.

The starting configurations for simulated liquid systems can be constructed in different ways. A common approach is to start from a known crystalline structure and melt the system by equilibrating at high temperature. Another option is to start from a random configuration and add constraints or restraint to the atomic motion during the equilibration phase of the simulation. Here we instead create a randomised starting geometry that is somewhat relaxed by geometric rather than dynamic means to avoid overlap of atoms (see Section 3.1). The initial velocities of the atoms are usually drawn from a Maxwell-Boltzmann distribution with the additional constraint that the sum of all momenta is zero to avoid drift of the whole system.

The differences between different MD methods lie in how these forces are calculated. In *ab initio* MD (AIMD) the electronic structure is computed using quantum chemical methods, while for classical MD, expressions for the interatomic forces need to be specified.

2.2.1 *Ab initio* Molecular Dynamics

Ab initio (lat. from first principles) in the context of MD simulations reflects that no parameters are needed to simulate the dynamics of the system. However, the most common method for evaluating forces is DFT (which is not considered *ab initio* in the context of quantum chemistry).

The most naive implementation of AIMD based on the TISE is the Born-Oppenheimer MD (BOMD) method [75]. It starts by initializing the atomic positions and momenta, whereafter the electron structure is computed by DFT. Based on the converged electron density and total energy, the forces acting on each atom are computed from the gradient of the energy upon perturbing the position of each atom in turn. The force on atom i is given by

$$\boldsymbol{F}_i = -\frac{\partial E}{\partial \boldsymbol{R}_i} \tag{2.12}$$

in the absence of magnetic or time-dependent external forces. Given the positions, momenta (and thereby velocities) and forces (and thereby accelerations) of all atoms, the system can be propagated to the next time-step where a new electron density needs to be calculated. However, this is typically much cheaper than the initialization since the first guess can be based on the converged electron structure from the previous time-step. The time-steps therefore need to be small enough to limit changes between consecutive snapshots in the electronic structure.

Even so, to update the electronic structure quantum mechanically for every snapshot is computationally expensive. Car and Parrinello therefore proposed

a method in 1985 based on a hybrid quantum/classical approach [91]. The problem is framed in terms of a total interaction energy functional dependent on both Kohn-Sham orbitals, ϕ_i, atomic positions, R_I and Lagrange multipliers α_ν, to fulfill any external constraints on the system (*e.g.* volume). The orbitals and Lagrange multipliers are assigned fictitious masses, μ and μ_ν in order to give them classical dynamics. This procedure gives the Lagrangian

$$L = \sum_i \frac{\mu}{2} \int_{\boldsymbol{R}^3} |\dot{\phi}_i|^2 d\boldsymbol{r} + \sum_I \frac{M_I R_I^2}{2} + \sum_\nu \frac{\mu_\nu \alpha_\nu^2}{2} - E[\{\phi_i\}, \{R_I\}, \{\alpha_\nu\}], \quad (2.13)$$

giving rise to the classical equations of motion:

$$\mu \ddot{\phi}_i(\boldsymbol{r}, t) = -\frac{\delta E}{\delta \phi_i^*(\boldsymbol{r}, t)} + \sum_k \Lambda_{ik} \phi_k(\boldsymbol{r}, t), \quad (2.14)$$

$$M \ddot{R}_I = -\nabla_{R_I} E, \quad (2.15)$$

$$\mu_\nu \ddot{\alpha_\nu} = -\frac{\partial E}{\partial \alpha_\nu}, \quad (2.16)$$

where Λ_{ik} are Lagrange multipliers to maintain orthonormality of Kohn-Sham orbitals.

The fictitious masses of the electrons need to be selected to be sufficiently lighter than the nuclear masses so that no energy is exchanged between these degrees of freedom in the simulated dynamics. However, the electron mass should be chosen much heavier than the physical value in order to allow longer simulation time-steps than possible in BOMD simulations. A typical choice is 400 times the electron rest mass, which we have used in papers I and IV.

2.2.2 Classical Molecular Dynamics

In classical MD the forces are computed classically, and therefore a simulation is computationally much cheaper than AIMD simulation, but since it is dependent on (a quite large number of) empirical parameters, the quality is very much dependent on the parameters used [76, 77].

In materials science, such a parameter set is usually known as an interatomic potential. Since the materials, mostly crystalline solids, often are quite simple with a small number of distinct interactions, relatively few parameters are needed.

Computational chemistry on the other hand, tends to deal with molecules, which are held together by a multitude of different bonds, whose character

and physical origins differ, *e.g.* covalent bonds of different orders, ionic bonds, hydrogen bonds, and additionally also van der Waals interactions. In order to describe all of these accurately by classical models, it is usually not enough to only consider which pairs of elements are bound, but also their respective local environments – a carbonyl oxygen atom is different from an ether oxygen atom, etc. This can lead to a very large number of parameters in practice. The set of potentials in computational chemistry are conventionally called *force fields* [77].

There is much experience in using classical MD to study organic chemistry based systems, which over time has resulted in more or less standardized functional forms for the terms used in force fields [77, 92, 93]:

$$E = 4\sum_{(i,j)} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{1}{4\pi\epsilon_0} \sum_{(i,j)} \frac{q_i q_j}{r_{ij}} + \sum_{\text{bonds}} \frac{k_b (l - l_{0,b})^2}{2} + \sum_{\text{bond angles}} \frac{\kappa_a (\theta - \theta_{0,a})^2}{2} + \sum_{\text{dihedrals}} \sum_{n} a_{n,d} \cos(n\phi + \phi_{0,d}), \tag{2.17}$$

where the first term is the Lennard-Jones approximation for the Pauli exclusion principle and van der Waals interactions, the second term is the Coulomb interactions between all pairs of atoms, the third term describes harmonic bond potentials, the fourth harmonic potential for bond angles and the last a Fourier series for the energy of proper and improper dihedrals, *i.e.*, torsions and out-of-plane bending of groups of four bonded atoms (Fig. 2.1).

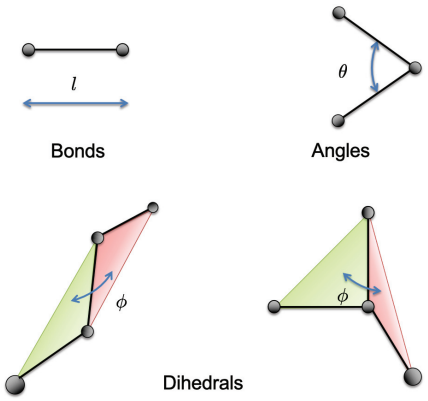


Figure 2.1: Schematic illustration of some interactions included in force fields for organic systems.

The parameters of a force field of the form (2.17) are thus a set:

$$\mathrm{FF} = \{\epsilon_i, \sigma_i, q_i, k_b, l_{0,b}, \kappa_a, \theta_{0,a}, a_{n,d}, \phi_{0,d}\}_{i,b,a,d}\,, \tag{2.18}$$

where $\sigma_{ij} = (\sigma_i + \sigma_j)/2$ and $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$.

The non-bonded, *i.e.* Lennard-Jones and Coulomb terms, are typically excluded for directly bonded atoms and next-nearest neighbours, and scaled down to a number between 0 and 1 for third neighbours.

The force field parameters are typically based on a combination of quantum chemical computations for single molecules in vacuum and experimental measurements, usually thermodynamic in nature. They are also usually validated by comparing their predictions of thermodynamic properties with experimentally observed values. This means that there is often no guarantee that a force field will accurately reproduce the microdynamics even if it is observed to give good fit with macrodynamics.

2.3 Graph Theory

Our analysis needs to represent the topological structure of physical systems by the manner in which atoms and molecules are connected. This is done by a graph representation detailed in Chapter 3. In this section, some general theory of graphs is presented to the extent needed to understand the representations of molecular structure within CHAMPION.

Graph theory is the branch of mathematics that deals with how things are connected. A graph is in general composed of two types of entities: *vertices* (or *nodes*) and *edges* (or *links*). A vertex is by itself a formless abstract entity with no data associated with it apart from a unique index. An edge connects two vertices. It is thus the edges that give the graph its structure. Edges can be either *undirected* or *directed*. For example, 'A causes B' could naturally be encoded by a directed edge from vertex A to vertex B. A bond between two atoms, on the other hand, is naturally represented by an undirected edge since the relation is symmetric.

An undirected graph could *e.g.* be represented by a symmetric matrix M of 1s and 0s where $M_{ij} = 1$ signifies that vertices i and j are connected by an edge, and $M_{ij} = 0$ that they are not. Since many real-world graphs are sparse (*i.e.* most pairs of vertices do not share an edge), it is often more practical to encode the graph in a sparse representation, *i.e.* as a list of pairs of vertex indices denoting the edges.

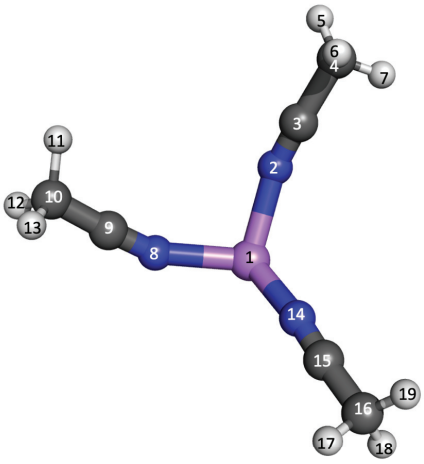
While a graph lacks any additional information about what it encodes, one typically has in mind an interpretation of what the vertices and edges

represent, that may have extra information associated with it. For example, in a bond graph each vertex represents an atom and each edge represents a bond between atoms. Atoms are of different chemical elements, and chemical bonds can have associated with them different lengths, stiffnesses and bond order *etc.* *Labeled graphs* generalise the concept of graphs to cases where vertices and edges can have different discrete *equivalence classes*, *i.e.* types, *e.g.* as a result of different sets of data associated with vertices and/or edges.

Two graphs (labeled or unlabeled) are *isomorphic* if they are identical up to a permutation of the vertex indices. intuitively, two isomorphic graphs have the same network structure. This makes isomorphism an important property to be able to assess. Unfortunately, it is also a hard problem in general since all commonly used computational representations of graphs are contingent on the indices assigned to the vertices, and the number of permutations grows as the factorial of the number of vertices of each equivalence class, rapidly becoming intractable. The most practical approach to this problem is by applying graph *canonisation*, *i.e.* indexing the vertices in a canonical order, such that any swap of vertex indices takes the graph to an uncanonical form unless the two vertices have indistinguishable positions in the graph, as well as identical labels in the case of a labeled graph. Testing whether two canonised graphs are isomorphic reduces to testing whether their representations are identical. While the computational complexity of graph canonisation is still an open question, there exist algorithms that tend to work well in practice, often in linear time on average.

It is often useful to partition large graphs into smaller subgraphs. One way of achieving this is to decompose a graph into as many disjunct subgraphs as possible without any edges connecting vertices in different subgraphs. Another way of putting it is to subdivide the graph as far as possible without cutting any edges. The subgraphs thus obtained are called *connected components*, or simply *components*. An example of a bond graph is shown in Fig. 2.2.

It is sometimes useful to classify the network topology locally around vertices of a specific equivalence class. One can then start from each such vertex and include the vertices and edges up to a predetermined number of edges removed from the starting point, *i.e.* the *graph distance*. Such a *graph embedding* is defined by the combination of its graph and its root vertex.



1	2
1	8
1	14
2	3
3	4
4	5
4	6
4	7
8	9
9	10
10	11
10	12
10	13
14	15
15	16
16	17
16	18
16	19

Figure 2.2: Example of a bond graph (left) and the graph canonised according to CHAMPION's rules (right).

2.4 Statistical Physics

Statistical physics is the discipline of applying statistics and probability theory to the physics of large numbers of interacting particles to predict their *ensemble* behaviour, *i.e.* their behaviour in the large number limit. Statistical physics thus connects two different scales, the *microscopic scale* and the *macroscopic scale*.

Often, but not always, the two scales tend to be described by different sets of variables. For example, the atoms and molecules in a liquid are most conveniently described by their positions and velocities as functions of time (*i.e.* their trajectories) while the liquid at the macroscopic scale is more readily described in terms of thermodynamic properties such as temperature, pressure, density and viscosity. Statistical physics leverages the tendency of random errors to cancel out, so that mean values of the macroscopic variables can be predicted based on the central limit theorem (CLT).

In order for statistical physics to map the properties on the microscopic scale to those on the macroscopic scale models are needed for how the macroscopic variables emerge from the interactions of the microscopic variables. Traditionally, these models were probabilistic models based on known laws

of physics integrated over all possible initial and boundary conditions. Since the advent of MD simulations, the statistics at the microscopic level can alternatively be "harvested" from the simulated behaviour of a system of atoms or molecules.

Using these models in their "forward" direction entails predicting the values of the macroscopic variables based on the microscopic dynamics. They can also be applied in the "backward" direction by testing hypotheses about the microscopic dynamics by matching their predictions against macroscopically observed variables. Examples of this approach include the triumphs Boltzmann and Maxwell in explaining classical thermodynamics by the kinetic theory of gases, essentially reproducing the known thermodynamics of the gas phase from a Newtonian "billiard ball" model of atoms.

Statistical physics universally relies on the law of large numbers, which is a very reasonable assumption for real macroscopic physical systems, on the order of 10^{23} atoms. For simulated systems, one has to be more careful since the number of simulated atoms is more often on the order of thousands or even just hundreds of atoms. By an additional assumption, that of *ergodicity*, ensemble averages, which intrinsically depend on very large numbers of atoms, can be complemented by time averages of fewer atoms. While ergodicity of dynamical systems is hard to prove in general, this assumption seems to be reasonable for most real condensed matter systems for reasons that are not entirely understood. Thus one must then take care to not overweight the values of nearby points in time since most variables have a correlation time that needs to be taken into account. Statistical physics does, however, have the advantage that within the applied assumptions, the uncertainties of all predictions can be quantified.

2.4.1 Coordination and Solvation Numbers

Coordination number (CN) and solvation number (SN) are both measures for quantifying the composition of the first solvation shell. The CN is the number of coordination bonds of *e.g.* a Li^+ ion, whereas the SN is the number of ligands. Thus CN and SN are often the same, but differ whenever two or more atoms of the same solvent or anion are coordinated to the central Li^+ ion. To further quantify the 1^{st} solvation shell compositions, one can further break the CN down into the partial CN (pCN) of each coordinated species, such they add up to the total CN, and similarly for SN (pSN).

Figure 2.3 shows an example Li^+ 1^{st} solvation shell. Here the SN is 3 because the Li^+ ion has 3 ligands, and the CN is 4 due to the 4 coordination

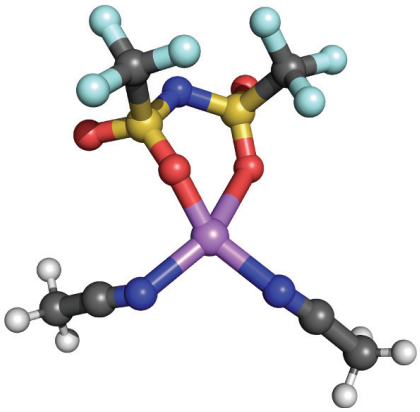


Figure 2.3: Example of Li^+ 1st solvation shell.

bonds. The pCN is 2 for the sulfonyl oxygen atoms and the nitrile nitrogen atoms and the pSN is 1 for TFSI and 2 for ACN.

We finally introduce the notion of pSN as a measure of the effective interaction strength of different coordination bonds. The pSN overrepresentation is given by the fraction of different species in the first solvation shell compared to what would be expected from purely stoichiometric considerations:

$$O = \frac{P_i}{P_{0i}} - 1, \tag{2.19}$$

where P_i is the fraction of SN attributed to pSN_i, and P_{0i} is species i's stoichiometric fraction of ligands.

2.4.2 Species Diffusivity

An example of a phenomenon often treated by statistical physics, of large relevance for this work, is diffusion. From a phenomenological point of view diffusion is the process by which concentration differences decrease and equilibrate over time, as can *e.g.* be seen when a drop of dye is dripped into a glass of water, at first dispersing chaotically, to eventually becoming evenly spread out throughout the glass. In a battery cell, it can be used to describe the thermodynamic drive towards eliminating salt concentration gradients caused by the electrostatic fields near the electrode surfaces, but is also connected to transport processes more broadly in the electrolyte, also in the absence of concentration gradients.

The microscopic understanding of diffusion is based on Einstein's 1905 treatment of Brownian motion [94]. While the original theory was developed

for large particles in a medium of much smaller ones, because of the CLT, the aggregate effect over time scales covering many collisions is the same also for particles of similar size to those in their environment, and this formalism is ubiquitously used in simulations of liquids [66]. Consider a liquid consisting of particles, *e.g.* ions and molecules, in which we single out a single particle. The particles in the liquid have velocities on the thermal scale: $v \approx \sqrt{\frac{3k_BT}{m}}$, around 1000 m/s for a Li^+ ion. According to Newton's first law, the particle's velocity is maintained in the absence of interactions with other particles. However, as interactions are ubiquitous in condensed matter, the velocity will be chaotically changed as the particle interacts with its environment, leading to a gradual "loss of memory" w.r.t. the initial velocity.

To model the diffusion on the microscale, we imagine that the particle's velocity de-correlates during a time interval of length τ. Since the motion within every interval is independent of the motion within every other interval while being generated by the same physics, the successive displacements are independent and identically distributed (i.i.d.), and additionally with finite mean and variance, which means that the CLT holds. We can therefore simplify the dynamics by considering each displacement to be of equal length $l = \langle \Delta r(t, t+\tau) \rangle$. To further simplify the model we consider the motion taking place on a lattice rather than in continuous space, so that at each step the motion is parallel (or antiparallel) to one of the coordinate axes. This simplifies the analysis while also not affecting the macroscopic results. We also assume that the liquid is isotropic, so that the motion is along each cardinal direction with probability 1/6. We first consider only the motion in the x direction.

The probability of the particle being found at x at time $t + \tau$ can be expressed using the law of total probability for its whereabouts at time t,

$$P(x, t + \tau) = \frac{1}{6}P(x - l, t) + \frac{1}{6}P(x + l, t) + \frac{4}{6}P(x, t). \tag{2.20}$$

This equation is Taylor expanded to the leading non-vanishing order,

$$\begin{aligned} P(x,t) + \tau\frac{\partial P(x,t)}{\partial t} + \ldots = \\ \frac{1}{6}\Bigg(& P(x,t) - l\frac{\partial P(x,t)}{\partial x} + \frac{l^2}{2}\frac{\partial^2 P(x,t)}{\partial x^2} + \ldots \\ & + P(x,t) + l\frac{\partial P(x,t)}{\partial x} + \frac{l^2}{2}\frac{\partial^2 P(x,t)}{\partial x^2} + \ldots \\ & + 4P(x,t)\Bigg), \end{aligned} \tag{2.21}$$

which after cancelling terms, dividing with τ and eliminating higher order terms simplifies to the partial differential equation

$$\frac{\partial P(x,t)}{\partial t} = \frac{l^2}{6\tau}\frac{\partial^2 P(x,t)}{\partial x^2}. \tag{2.22}$$

The same equation can be straight-forwardly adapted for concentrations of a species rather than for the probability distribution of a single particle, and generalised to 3-dimensional motion:

$$\frac{\partial c(\boldsymbol{r},t)}{\partial t} = D\nabla^2 c(\boldsymbol{r},t), \tag{2.23}$$

where c is the concentration, $\boldsymbol{r} = (x,y,z)$ and the dffusivity, $D = l^2/6\tau$. This is the well-known diffusion equation. Macroscopically, D can be measured by considering the speed of particle transport in the absence of forces biasing the transport in any particular direction.

It remains to be shown how to connect this result in terms of MD simulations of the microscale, where probability distributions and concentrations are not directly measurable. One way is by the method of mean squared displacement (MSD). Based on the model for the microdynamics, we can compute statistical moments of the position of the particle as a function of time. The mean displacement over a time $t = n\tau$ is

$$\langle \Delta r(0,t) \rangle = \left\langle \sum_{i=1}^{n} \delta \boldsymbol{r}_i \right\rangle = \sum_{i=1}^{n} \langle \delta \boldsymbol{r}_i \rangle = 0, \tag{2.24}$$

i.e. the most likely place to find the particle at any time is where it began.

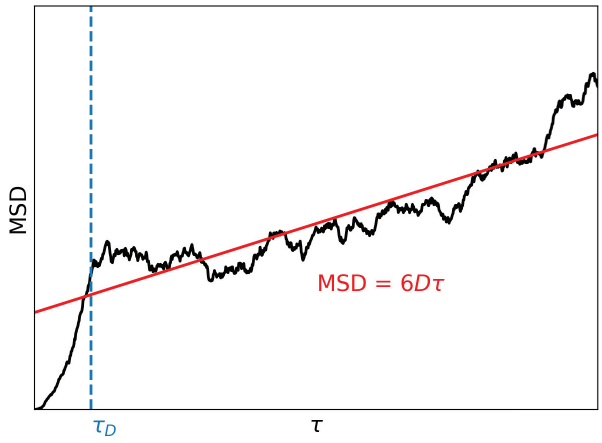


Figure 2.4: Schematic MSD curve, showing also τ_D, the diffusion onset time and the linear fit used to compute D.

Considering instead the MSD, we get

$$\left\langle \Delta \boldsymbol{r}(0,t)^2 \right\rangle = \left\langle \left(\sum_{i=1}^{n} \delta \boldsymbol{r}_i \right)^2 \right\rangle = \sum_{i,j=1}^{n} \langle \delta \boldsymbol{r}_i \cdot \delta \boldsymbol{r}_j \rangle = \sum_{i=1}^{n} \langle \delta \boldsymbol{r}_i^2 \rangle = nl^2. \quad (2.25)$$

Using further that $n = t/\tau$, we get

$$\left\langle \Delta \boldsymbol{r}(0,t)^2 \right\rangle = \frac{l^2}{\tau} t = 6Dt, \quad (2.26)$$

using our earlier result for the dffusivity. This directly relates the dffusivity with the MSD, which can be directly computed as an average over all particles of a given species, using all time points as starting points in the average:

$$\left\langle \Delta \boldsymbol{r}^2 \right\rangle (\tau) = \left\langle \Delta \boldsymbol{r}(t, t+\tau)^2 \right\rangle_{k,t}. \quad (2.27)$$

This can be plotted as curve (Fig. 2.4), where 1/6 of the slope after the diffusion onset time τ_D gives the dffusivity.

2.5 Optimisation Methods

Optimisation has been used in this work to a) generate starting geometries for MD simulations, b) match orientations of molecules between different snapshots, c) fit curves, and d) fit the parameters of a force field.

Optimisation is the process of finding a minimum value of a *cost function* (a.k.a.*loss function*), here denoted by $C(\boldsymbol{x})$, associated with a given point $\boldsymbol{x}$ in *design space*, the space of independent coordinates that may be chosen, and where each point is assigned a cost value. An example would be an actual landscape where the optimisation problem is to find the lowest point. Design space is then the real 2-dimensional space along which you can move, and the cost function is the height above some arbitrary level (*e.g.* sea level).

Any found minimum might be local or global, and in general it is not possible to know which, but this is also strongly method dependent. Genetic algorithms (GAs, see 2.5.2) are designed specifically to avoid getting stuck in local minima, while for example most gradient-based methods (2.5.1) have a very strong tendency to only locate the closest local minimum. Yet, when a local optimum suffices, or the solution domain is known to be convex, gradient-based methods are practical due to their simplicity.

2.5.1 Gradient-Based Optimisation

Gradient-based optimisation is a class of (usually) deterministic methods used to find the closest local minimum by taking into account the slope, and possibly higher order derivatives of the current best guess in order to select the next point in design space to evaluate. These methods work best for relatively low-dimensional design spaces and a requirement is that it is possible to evaluate the gradients

$$\nabla C = \frac{\partial C}{\partial \boldsymbol{x}} \tag{2.28}$$

of the cost function with respect to small movements in design space. This may be done analytically or numerically by finite differences – for the latter it is intuitive from Eq. (2.28) that the computational cost is proportional to the number of dimensions along which $\boldsymbol{x}$ needs to be perturbed.

Steepest descent is the most straight-forward gradient-based optimisation method and always moves in the direction in which the cost function declines most steeply, *i.e.*, along the negative gradient. The step length is determined by a line search in the search direction, which finds the minimum using one-dimensional convex methods, such as Newton's method:

$$\Delta x = -\frac{f'(x)}{f''(x)}. \tag{2.29}$$

While the steepest descent method is simple to both understand and implement, it has a tendency to zigzag its way to the solution as the next gradient

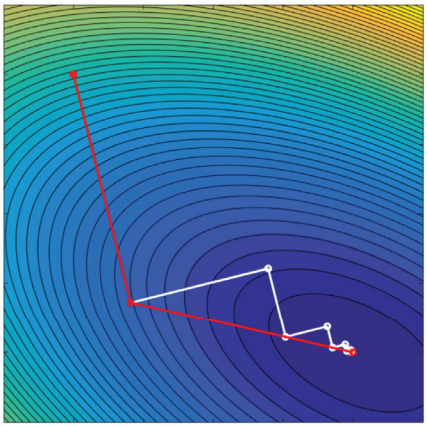


Figure 2.5: Schematic showing the convergence behaviour of steepest descent (white) and conjugate gradient (red) methods. [95]

will always be orthogonal to the previous search direction, making convergence suboptimal. This can be remedied by taking into account previous search directions; the conjugate gradient method, where subsequent directions are also conjugated with previous search directions based on the shape of the cost function landscape (Fig. 2.5). Conjugate gradient optimisation is in this thesis work used for geometry generation, orientation matching and curve fitting.

2.5.2 Genetic Algorithms

Genetic algorithms (GAs), a.k.a. evolutionary algorithms, are a family of stochastic optimisation heuristics based on simulating the process of Darwinian evolution [96]. The objective is for successful solutions to arise over many iterations of an evolutionary process. The aim is to over many generations find a successful *genotype*, *e.g.* a set of force field parameters, that produces the fittest *phenotypes*, *e.g.* reproducing atomic forces accurately.

The genotypes need to be encoded in some manner, and in this work real-valued encoding was used. In addition to being relatively resistant against local minima, GAs are practical when evaluating gradients is either computationally costly or the cost function is not smooth.

The basic structure of a GA is presented in Algorithm 1; A population is generated randomly with broad starting guesses for all parameters reflecting some knowledge of the reasonable values. They are thereafter evaluated, which underlies the selection of parents for the next generation, out of a softmax distribution (analogous to a Boltzmann distribution) premiering low-cost

solutions:

$$P_i = \frac{e^{-\beta C_i}}{\sum_j e^{-\beta C_j}}. \tag{2.30}$$

For each new genotype to be constructed, two parents are selected uniformly from the list of parents and recombined. Each new genotype is mutated, *i.e.*, its parameters are perturbed. The best few individuals survive unaltered to the next generation.

Algorithm 1: Basic structure of a GA

```
Initialize a population of individuals from broad starting guesses;
while not converged do
    Evaluate costs C_i;
    for i = 1, i ≤ number of elites, i++ do
        Save the ith best individual to next generation
    end
    for each parental slot do
        Select parent, P_i = e^{-βC_i} / Σ_j e^{-βC_j};
    end
    for each new individual do
        Choose parent a;
        Choose parent b;
        Cross-over(a, b);
        Mutate;
    end
end
```

Chapter 3

Developed Methods

The ultimate vision guiding this work has been to enable computational modelling to elucidate the properties and behaviour of complex condensed matter systems on all relevant scales, with a focus on tackling the challenges of modelling battery electrolytes (Section 1.5). While this is a much too tall order to fit into a single PhD project, I have strived to build a good bottom-up foundation that can be further built upon, and extended by complementary approaches.

This framework (Fig. 3.1) is strictly bottom-up to put it on firm ground and avoid ad-hoc parameters. Properties predicted at one level are analysed to find patterns to be aggregated and used at the next level. The ground level is AIMD. Simulations on this level can currently cover hundreds of atoms over picoseconds at a reasonable computational cost. This is sometimes not enough to give an accurate description of the bulk phase of the simulated system. It could then instead be used to fit a classical force field to reproduce the forces on atomic nuclei according to the AIMD simulation by machine learning. Using a physics based force field rather than a pure machine learning approach, such as neural networks, should reduce the needed amount of training data considerably due to the smaller number of parameters. In contrast to conventional transferable force fields, these would be tailor-made for each system with minimal need for expert knowledge and manual labour. With classical MD, simulations on the scale of thousands of atoms over nanoseconds is practically feasible. For simple liquids made up of weakly interacting molecules, standard statistical physics may suffice to model and explain the bulk behaviour. For more complex liquids forming dynamic supramolecular structures, however, more elaborate analysis is required, as discussed in Section 1.5, that can recognise these structures and subsequently characterise their structure, dynamics and transport properties. While outside the scope of this thesis, the results from this analysis can further be straightforwardly input into *e.g.* a kinetic Monte Carlo (kMC) method, or even continuum population balance and transport models. In addition to this MSM framework, I have also devel-

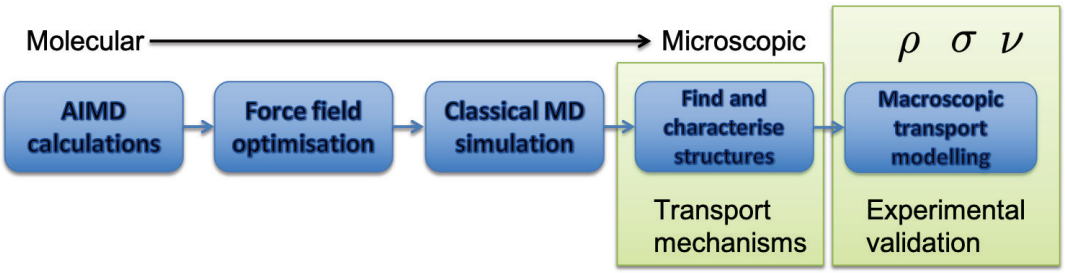


Figure 3.1: Overview of our envisioned MSM framework.

oped a system builder, that generates random initial configurations for MD simulations, that avoids overlap between atoms.

The following sections describe in turn, the starting geometry builder, the force field optimisation algorithm, and the DSD and analysis.

3.1 Starting Geometry Generation

The starting geometries for AIMD simulations are generated in two steps: first a randomised molecular geometry with the desired density is spread out at random in a periodic box. Thereafter, the configuration is relaxed iteratively by by minimising the cost function

$$C_i = \sum_{i \neq j} H\left(R_i + R_j - d_{ij}\right) \left(1 - \frac{d_{ij}}{R_i + R_j}\right)^2, \tag{3.1}$$

where the sum is over all atoms, d_{ij} denotes the distance between atoms i and j, and R_i is the van der Waals radius of atom i. The cost is minimised w.r.t. translation and rotation of each molecule or ion in turn, using the conjugate gradient method to find the closest minimum given the present positions of all other molecules.

3.2 Force Field optimisation

The functional form of a system-specific force field is selected in advance, whereafter the parameters of the force field are optimised by GA. A population of individuals is generated from normal distributions with variances proportional to the *a priori* uncertainty in the correct values. The number of individuals should be at least on the same order as the number of parameters,

and the number of parents should be approximately a third of the population size, based on trial and error. Individuals are assessed based on their ability to reproduce the forces predicted by AIMD. In each training step a small number of snapshots from the AIMD trajectory are selected and the forces on each atom computed using the trial force fields. The choice of using the forces for training, rather than the energy, is motivated by the fact that the dynamics is uniquely determined by the forces, and that there is a force vector acting on each atom in contrast to just one global value of the energy.

The cost function used is the sum of a direction cost and a magnitude cost,

$$C = \underbrace{A \tan\left(\frac{\pi}{4}\left(1 - \hat{\boldsymbol{F}}_{\mathrm{ref}} \cdot \hat{\boldsymbol{F}}_{\mathrm{FF}}\right)\right)}_{\text{Direction cost}} + \underbrace{B \log^2\left(\frac{\boldsymbol{F}_{\mathrm{ref}}^2}{\boldsymbol{F}_{\mathrm{FF}}^2}\right)}_{\text{Magnitude cost}}, \tag{3.2}$$

where $\boldsymbol{F}_{\mathrm{ref}}$ and $\boldsymbol{F}_{\mathrm{FF}}$ are the AIMD and force field forces, respectively, $\hat{\boldsymbol{F}}_{\mathrm{ref}}$ and $\hat{\boldsymbol{F}}_{\mathrm{FF}}$ are the corresponding unit vectors and A and B are parameters setting the relative importance of each term. The direction cost is constructed so that it gives a cost of 0 to parallel force vectors and approaches infinity for anti-parallel forces, while the magnitude cost function is 0 if the magnitudes are equal and grows symmetrically in the order of magnitude discrepancy between $\boldsymbol{F}_{\mathrm{ref}}$ and $\boldsymbol{F}_{\mathrm{FF}}$. Such a partition of the cost function was found necessary to avoid the tendency of force constants to decay to 0 early in the optimisation. In order to track the progress of the optimisation, the relative force error, defined by

$$E_{\mathrm{rel}} = \left\langle \frac{|\boldsymbol{F}_{\mathrm{FF}} - \boldsymbol{F}_{\mathrm{ref}}|}{|\boldsymbol{F}_{\mathrm{ref}}|} \right\rangle, \tag{3.3}$$

where the average is over evaluated forces, is used as a more intuitive measure of error.

Most of the production data is used in training the force field, but a small part, *ca.* 1000 snapshots, is reserved for out-of-sample testing, to ensure that the learned force field is not only optimised for the specific configurations (snapshots) on which it has been trained (*i.e.* overfitting). The training and testing data should preferably be uncorrelated, which is most easily achieved by discarding part of the trajectory between the training and testing data. The training is successfully completed when the relative force error on the testing data is lower than the chosen tolerance.

3.3 Dynamic Structure Discovery

The DSD method is based on finding and analysing what moves together. It automatically discovers time-dependent bonds between atoms, classifies the structures defined by these bonds using graph theory, and computes the physicochemical properties of the structures using statistical physics.

3.3.1 Bond detection

The first step of the DSD method, (pat. pend. [97]) is to detect the bonds between atoms in the simulation trajectory, and if applicable their time of formation and breakage. The idea behind the method is to find what moves together by first identifying candidate bonds as pairs of atoms that are relatively close together over a period of time, and then subjecting the candidate bonds to a series of tests.

Candidate bonds are identified as pairs of atoms that are closer than a scaled sum of their atomic radii, *e.g.* 0.8 times the sum of their van der Waals radii (Fig. 3.2a). The candidate bond is considered only during the time period when this condition is fulfilled. Over this time period, an average distance between the atoms is computed. If the atoms are indeed bound, one would expect their average distance over the lifetime of the bond to be very close to the equilibrium length for bonds between atoms of their respective types. This equilibrium bond length should also coincide with the most probable distance between the atomic types, which can be obtained *e.g.* by computing their partial radial distribution function (pRDF),

$$g_{ij}(r) = \frac{1}{n_0} \frac{n(r)}{4\pi r^2}, \tag{3.4}$$

where $n(r)$ is the number density of neighbours of atom type j on distance r from atoms of type i and the expression is normalised by the average bulk number density n_0 of species j. The equilibrium bond length, if the trajectory contains bonds between atom types i and j, should be the first peak in this function (Fig. 3.2b). The second test of a candidate bond is thus whether the average distance between the atoms is within a tolerance of this peak:

$$(1-\alpha)\, r_{\text{peak}} \leq \frac{1}{T} \int_t^{t+T} d_{ij}(t)\, dt \leq (1+\alpha)\, r_{\text{peak}}. \tag{3.5}$$

The tolerance α is itself proportional to the width of the pRDF peak, as assessed by *e.g.* the full width at half maximum (FWHM).

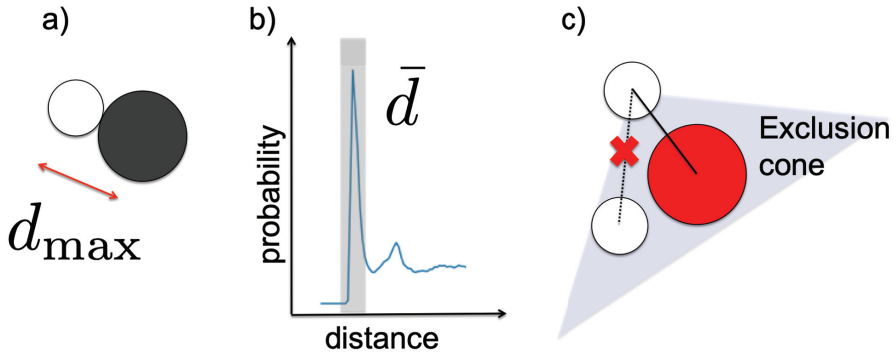


Figure 3.2: Criteria for whether a pair of atoms are bound.

The third and final requirement on a candidate bond is designed to eliminate false positives. For example, the two hydrogen atoms in a water molecule may well fulfill the above requirements (Fig. 3.2a-b). We know, however, that they are not truly bound, but that the reason why they are relatively close to each other, and why their average distance coincides with the first peak of their pRDF, is that they are oscillating about this sharply defined distance because they are both bound to the same oxygen atom. How can we tell true bonds from spurious ones? A first attempt could be to eliminate the candidate bond between the hydrogens since it is longer than the bonds between each hydrogen atom and the oxygen atom. However, this does not work, as can be seen by considering which bonds to accept from the point of view of the oxygen atom. At any given time, one of hydrogen atoms may well be closer to the oxygen atom than the other, but we know that the oxygen atom is bound to both. A clue to the solution is that the angle between the hydrogens from the point of view of the oxygen atom is rather large (104.5°), while the angle between the oxygen atom and the other hydrogen atom from the point of view of a hydrogen atom is much smaller (37.8°). Indeed, it is a general feature of chemistry that different bonds involving a single atom tend to be rather well spread out across the unit sphere. Since very few atoms have more than eight bonds, few bonds should be within *ca.* 70° of one another. We can thus formulate an exclusion criterion for candidate bonds by combining their relative lengths with the angle between them, and demanding that a candidate bond involving an atom is not within the exclusion cone subtended by a chosen angle about a shorter bond involving the same atom (Fig. 3.2c) if the ratio between the longer and the shorter bond is above a certain maximum allowed value (which for most systems can be set to 1).

If a candidate bond is accepted, its time of formation and breakage are subsequently determined. The distance between the atoms in a bond tend to oscillate sinusoidally with a certain amplitude about the equilibrium bond length. This amplitude is measured, and the bond formation time is set to the first time the interatomic distance goes below the furthest distance during oscillation (*i.e.* the equilibrium bond length plus the amplitude). The breakage time is correspondingly set to the last time the distance goes above this value.

3.3.2 Structure Classification

The bond detection algorithm gives all the information needed to construct a time-dependent global bond graph that uniquely determines the bond topology of the entire system. This bond graph is an undirected labeled graph, where the vertices represent the atoms, and are labeled either by the unique indices of each atom, or by the type of each atom. The edges, on the other hand, are unlabeled since the bond detection algorithm does not distinguish bond orders (Fig. 2.2).

The global bond graph can be subdivided into canonised subgraphs along the lines described in Section 2.3, into either components (*e.g.* molecules or ionic aggregates) or solvation shells (*e.g.* solvation shells). These subgraphs are subsequently classified by their bond graph topology, enabling different exemplars of the same type to be recognised, so that they can be analysed by statistical physics methods.

3.3.3 Characterisation of structure, dynamics and transport

Once all unique structures of interest have been identified, they can be analysed using any methods of statistical physics defined for a many-body system, including their geometries, masses, volumes, moments and products of inertia, populations, lifetimes, transitions rates, diffusivities *etc.*

While not all of these properties have been presented in the included papers, or implemented into code, doing so would be a rather straight-forward exercise in applying well-known equations from statistical physics [66]. For those that have been implemented, the details can be found in paper III. However, an additional innovation from this work relates to how to compute the contributions from different structures, transitions and modes of motion to the transport properties of a given species, applicable even if the structures are short-lived.

Considering a Li^+ ion, we would like to be able to decompose its total diffusivity into additive contributions. The first step is split the instantaneous velocity into parts from translation of the local structure, rigid body rotation of the local structure, and the relative velocity of the Li^+ ion relative to the local structure. The first of these contributes to the vehicular transport mechansim, the other two to non-vehicular mechanisms. By introducing indicator functions χ_i for when the ion is in structure type i and ξ_{ij} for when the transition closest in time is from type j to i, the velocity is decomposed into

$$\boldsymbol{v}(t) = \sum_{i,s} \chi_i \boldsymbol{v}_{\text{vehicular}}(t) + \sum_{ij,s} \xi_{ij} \left(\boldsymbol{v}(t)_{\text{structural}} + \boldsymbol{v}_{\text{rotational}}(t) \right) \tag{3.6}$$

or

$$\boldsymbol{v} = \sum_k \boldsymbol{v}_k \tag{3.7}$$

for short, where k represents all the free variables or any combination of them and the time coordinate has been made implicit. The MSD is given by

$$\text{MSD}(\tau) = \left\langle \Delta \boldsymbol{r}(t, t+\tau)^2 \right\rangle \tag{3.8}$$

where the average goes over all exemplars and starting points t. Its derivative is

$$\frac{d\text{MSD}}{d\tau} = 2 \left\langle \Delta \boldsymbol{r}(t, t+\tau) \cdot \boldsymbol{v}(t+\tau) \right\rangle, \tag{3.9}$$

where $\boldsymbol{v}(t+\tau)$ can now be expanded into its constituents

$$\frac{d\text{MSD}}{d\tau} = 2 \sum_k \left\langle \Delta \boldsymbol{r}(t, t+\tau) \cdot \boldsymbol{v}_k(t+\tau) \right\rangle, \tag{3.10}$$

and finally

$$D_k^* = \frac{1}{3} \left\langle \Delta \boldsymbol{r}\left(t, t+\tau\right) \cdot \boldsymbol{v}_k\left(t+\tau\right) \right\rangle \tag{3.11}$$

is the additive decomposition. See paper III for a fuller motivation and derivation of this result.

Chapter 4

Results

We have studied a number of quite different battery electrolytes using the common framework of CHAMPION analysis. Here the electrolyte concepts and the properties targeted are briefly summarised. In papers I and IV, we studied LiTFSI in ACN in different molar ratios to understand the structure, dynamics and ion transport mechanisms specific to HCEs, as well as the transition from conventional electrolytes to HCEs. Paper I also served as the proof-of-concept of the DSD method and its implementation in CHAMPION for a single HCE. Paper IV used four different salt concentrations to elucidate trends and the transition from conventional to highly concentrated behaviour.

In paper II, a mixing of two salts in a chelating solvent was found to, at appropriate ratios, give rise to ion transport synergies based on the interacting natures of the two salts.

Paper V is a little different as it does not treat a novel system, like HCEs, but the most well-known electrolyte in LIBs, LP30, the progenitor of most electrolytes used in commercial LIB cells today. This electrolyte has often been characterised as a simple dilute electrolyte but here we give a detailed account of the global and local structure, solvation dynamics and species transport mechanisms.

Paper III is again quite different as it is a method focused paper that gives the most detailed description of the theory behind DSD and CHAMPION, while not presenting any new analyses of physical systems.

In the following we use mainly papers I-II and IV-V, but also some previously unpublished results from the same studies, to present key findings by topic, starting with the overall structures of the electrolytes, followed by local structure in terms of Li^+ 1^{st} solvation shells, the structural dynamics of these solvation shells, and finally species transport and mechanisms with special focus on Li^+ ion transport.

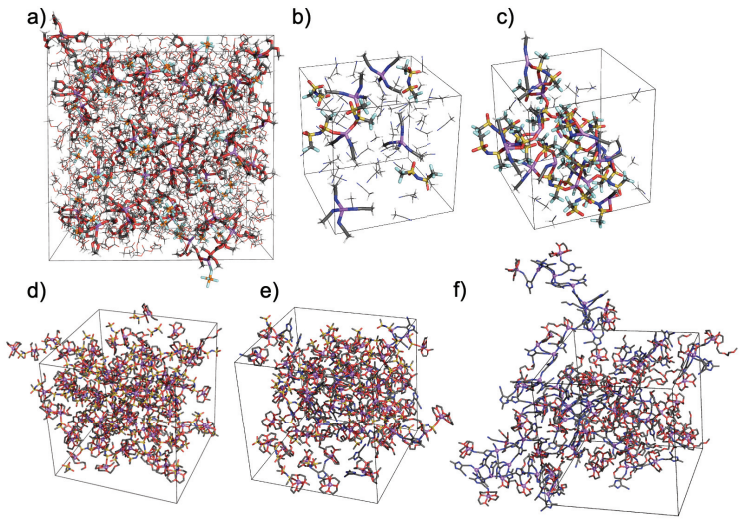


Figure 4.1: Global electrolyte structures, with thick bonds shown for ion-containing connected components: a) LP30, b) low concentrated LiTFSI in ACN, c) HCE LiTFSI in ACN, d) HCE LiTFSI in G4, e) HCE hybrid LiTFSI/LiTDI in G4, f) HCE LiTDI in G4. The bounding box represents PBCs and only one image of each atom is shown. H and F are hidden in d-f for greater visibility. Element colors: H: white, Li: purple, C: grey, N: blue, O: red, F: cyan, S: yellow.

4.1 Global Structure

By finding all the time-dependent bonds in a simulated system, CHAMPION can build its global bond graph by DSD. Visualising this bond graph by drawing the atoms and the bonds between them gives a good intuitive overview of the overall structure (Fig. 4.1).

We generally find that the degree of cation-anion interaction and aggregation increases with salt concentration, as expected, and consistent with the literature [35, 42, 72, 73]. LP30 is completely dominated by uncoordinated anions with only a small contribution from ion pairs (Fig. 4.1a), as is LiTFSI in ACN at low concentration (Fig. 4.1b).

For higher concentrations, gradually ion pairs and eventually aggregates become progressively more common. In the HCE regime, the nature of the global structure depends a lot on the exact balance of strengths between cation-anion and solvent interactions, and the intrinsic properties of the constituents; HCE LiTFSI in ACN forms a percolating network structure [35] (Fig. 4.1c),

as does HCE LiTDI in G4 (Fig. 4.1f). On the other hand, HCE LiTFSI in G4 (Fig. 4.1d) as well as the same electrolyte with a minor fraction of the salt replaced by LiTDI (Fig. 4.1e) form aggregates of several ions but not a percolating network.

In the case of LiTFSI in ACN, the flexibility of the TFSI anion [98] and the monodentate nature and small steric hindrance of Li-ACN coordination enables the formation of a percolating network with most but not all ACN molecules participating in Li^+ solvation without blocking access to Li^+ for TFSI (Fig. 4.1c). When LiTFSI is instead dissolved in G4, the chelating solvent blocks TFSI acccess to Li^+ much more effectively. While still usually allowing a Li^+ ion to coordinate two TFSI ions, they are constrained to be on opposite sides of the Li^+ ion, which constrains coordination too much to accommodate network formation. Instead, chains of alternating ions are formed, which terminate after a few links (Fig. 4.1d). The balance of interactions favours Li-G4 solvation over Li-TFSI interaction, leading to SIL structure. Still, the solvated Li^+ ions almost always coordinate one or two TFSI anions in addition to the G4 molecule wrapped around them.

The nature of network formation in LiTDI in G4 is entirely different. While the binding energy of a single ion pair is greater for LiTFSI than for LiTDI in vacuum both with and without the Li^+ ion being solvated, this result relies on bidentate LiTFSI coordination, whereas HCE composition favours TFSI ions to coordinate to two Li^+ ions each in the condensed phase, which steric considerations most often do not allow to be bidentate [39]. This shifts the relative balance between cation-anion *vs.* cation-solvent interaction for the two electrolytes, favouring aggregation of LiTDI over LiTFSI.

The rigidity of the TDI anion [99] and the greater separation of its four coordination sites compared with TFSI enables a rigid and disordered network to form where many of the coordination sites are vacant, giving rise to a spacious anionic network structure. The cations and solvent molecules not participating in this network instead forms a phase where Li^+ ions and G4 molecules intermingle polygamously in the absence of anions, thus not at all conforming to SIL structure [100]. Thus, the global structure is one of disproportionated micro-domains with very different properties (Fig. 4.1f). The hybrid electrolyte, while structurally intermediate to the pure LiTFSI and LiTDI electrolytes, does not have enough TDI for a percolating network to be formed (Fig. 4.1e).

Finally, we find that the common claim of HCE behaviour being determined by the complete absence of free solvent [32, 101], does not seem to be true in general, since for HCE LiTFSI in ACN, a substantial fraction of the

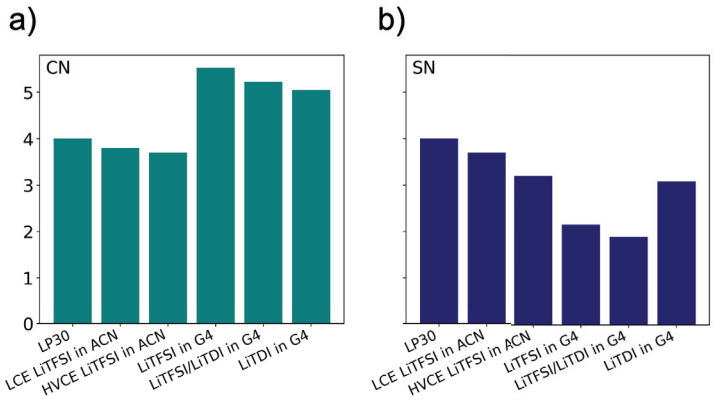


Figure 4.2: a) CN and b) SN for the different electrolytes.

ACN molecules were still found to be free (Fig. 4.1c). Instead, we propose that the lack of a continuous solvent phase and the predominance of coordinated anions combine to give the altered stability properties peculiar to HCEs. This is consistent with the AIMD and XPS results of Yamada *et al.* [35]. Chelating solvents, on the other hand, are very rarely completely uncoordinated to the cations whether or not they form SILs, due to their flexibility and large number of coordinating sites, all of which are sterically permitted to orient towards one and the same cation (Fig. 4.1d-f) [39].

4.2 Local Structure

By considering subgraphs of the global bond graph in an electrolyte, we analyse the statistics on the local structure around a central Li^+ ion, *i.e.* the 1^{st} solvation shell. We first consider the overall SNs and CNs, and then expand our analysis to consider the contributions from individual species, *i.e.* pSNs and pCNs. The pSNs should stand in proportion to the concentration of each species as well as their interaction strength with Li^+, but here again, steric effects also need to be considered; the effective interaction energies in the condensed phase may not be identical to those of structures optimised *in vacuo*.

The total CN for Li^+ is usually around 4 for electrolytes based on small organic molecules, such as ACN, DMC and EC, almost regardless of the composition [5, 102]. This is also what we find for LP30, as well as for LiTFSI in ACN regardless of salt concentration (4.2a). In many electrolytes, CN $\approx$

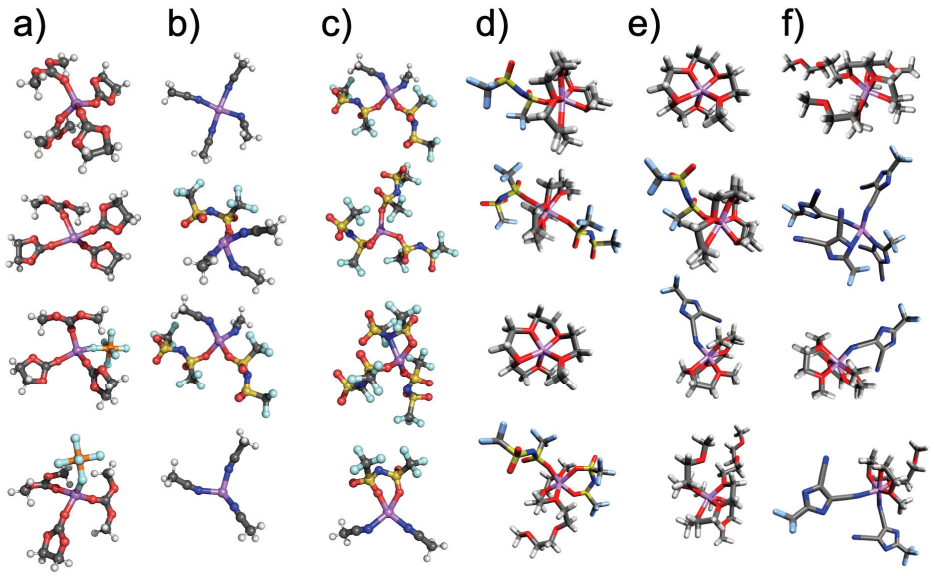


Figure 4.3: The 4 most common solvation shell structures of a) LP30, b) low concentration LiTFSI in ACN: c) HCE LiTFSI in ACN, d) HCE LiTFSI in G4, e) HCE LiTFSI/LiTDI (70/30) in G4, f) HCE LiTDI in G4.

SN, since most or all of the coordinations are monodentate. This is the case for LP30, but not for HCE LiTFSI in ACN, where the SN decreases from *ca.* 4 almost down to 3, due to the steric constraints of network formation in the HCE (Fig. 4.2b). The CN is, however, maintained around 4 due to a greater fraction of bidentate LiTFSI coordination at higher concentration, all this in good agreement with the results of Seo *et al.* [41].

Electrolytes based on chelating solvents, such as G4, stand out by having considerably higher CN [5, 102]. This since they can wrap around a Li^+ ion, sometimes with all ether oxygen molecules coordinated to the cation while still allowing anions to be coordinated in addition to the solvent. The CN decreases from 5.5 for pure LiTFSI in G4 to 5.0 for pure LiTDI in G4 (4.2a) . The SN, on the other hand, shows the opposite trend, going from *ca.* 2 for both the pure LiTFSI and the hybrid LiTFSI/LiTDI electrolytes to 3 for the pure LiTDI electrolyte (4.2b). The differences are readily explained in terms of the changes in global structure; SIL structure leads to high CN and low SN, whereas the disproportionated structure leads to relatively lower CN, due to less tightly wrapped G4 chains, and higher SN, due to multiple G4 chains and anions coordinating to each single Li^+.

Moving on to the composition of the 1st solvation shell, to gain an intuitive view, we first consider the 4 most common structures in each electrolyte (Fig. 4.3). Clearly, for the least concentrated electrolytes, as generally accepted, solvent molecules dominate the solvation shell, with anions also contributing to a lesser extent (Fig. 4.3a-b).

For HCEs, the picture is again more diverse. For HCE LiTFSI in ACN, anions dominate the 1st solvation shell (Fig. 4.3c), while for the electrolytes with chelating solvent, that is less clear (Fig. 4.3d-f). For the pure LiTFSI (Fig. 4.3d) and the hybrid (Fig. 4.3e) electrolytes, there are quite similar numbers of solvent molecules and anions in the solvation shell. For the pure LiTDI electrolyte (Fig. 4.3f) the anions again dominate, as for HCE LiTFSI in ACN, reflecting the greater relative strength of cation-anion interactions compared to cation-solvent. Here the disproportionated structure is again clearly visible.

We finally consider the overrepresentation of anions and solvent species in the Li^+ 1st solvation shell (Fig. 4.4). For LP30 EC is overrepresented while DMC is underrepresented, as reported elsewhere (Fig. 4.4a). Our overrepresentation of EC is somewhat smaller than reported elsewhwere [103]. PF_6^- is also overrepresented in the Li^+ 1st solvation shell, and more so than EC, reflecting the greater interaction strength of this interaction. LiTFSI in ACN exhibits by far the strongest overrepresentation observed in any of the electrolytes. Throughout the concentration range, TFSI is strongly overrepresented, indicating a strong preference for Li-TFSI over Li-ACN coordination (Fig. 4.4b).

The patterns for the LiTFSI/LiTDI in G4 electrolytes is the most subtle, but fully consistent with the above discussion of the causes of the global structure. In both pure electrolytes, the respective anion is overrepresented while the solvent is underrepresented (Fig. 4.4). This imbalance is stronger for TDI than for TFSI, as expected for its greater interaction strength in the condensed HCE phase. In the mixed electrolyte, the solvent is overrepresented, as is TDI, while TFSI is underrepresented. This again reflects the greater stability of Li-TDI coordination bonds compared to Li-TFSI.

4.3 Structural Dynamics

We next turn to the structural dynamics. Since CHAMPION outputs time-dependent bonds including coordination bonds, we can analyse the stability of Li^+ 1st solvation shells in terms of the width of their lifetime distributions. A general finding is that long lifetimes and high probability are highly correlated.

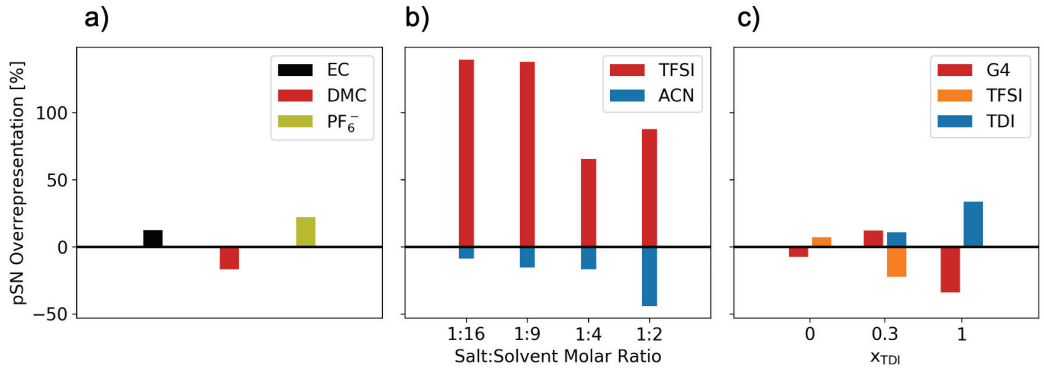


Figure 4.4: pSN overrepresentation for each solvent and anion species: a) LP30, b) LiTFSI in ACN, c) LiTFSI/LiTDI in G4.

This is quite natural, as longer lived species will naturally accumulate over time, whereas shorter-lived ones rapidly get converted into other species. For more detail, see Papers I, IV and V.

Another dynamic property of interest is the residence times of coordination bonds formed by Li^+ ions with solvents and anions. Mean and median residence times can be computed from bond survival curves, as was done in papers I, IV and V. Here the full survival curves give a clearer picture of the uncertainty and noise in the data (Fig. 4.5).

The vastly different time-scales involved in the structural dynamics show LP30 (Fig 4.5a) to be intermediate, with bonds surviving on the order of 100 ps. LiTFSI in ACN has two orders of magnitude faster dynamics, *ca.* 1 ps, and finally LiTFSI/LiTDI in G4 is at the other extreme, with coordination bonds lasting *ca.* 10 ns. The electrolytes thus span 4 orders of magnitude. Methodological differences (AIMD *vs.* classical MD) and non-equivalent parameter choices may well play a part here. On the other hand, the exponential dependence of the rate k on the energy barrier E_B in the Arrhenius equation, $k \propto e^{-k_B T/E_B}$, where k_B is the Boltzmann constant and T the temperature, makes such orders-of-magnitude differences plausible.

The quality of the data is also vastly different between the electrolytes. LP30 and the electrolytes based on G4 have been simulated by classical MD while LiTFSI in ACN has been simulated by AIMD. The shorter residence times overall for LP30 (Fig. 4.5a) compared to LiTFSI/LiTDI in G4 (Fig. 4.5c), gives many more transition events to sample, rendering the curves much smoother. For LiTFSI in ACN (Fig. 4.5b) the noise level is simply a consequence of much smaller simulation size and much shorter simulated time.

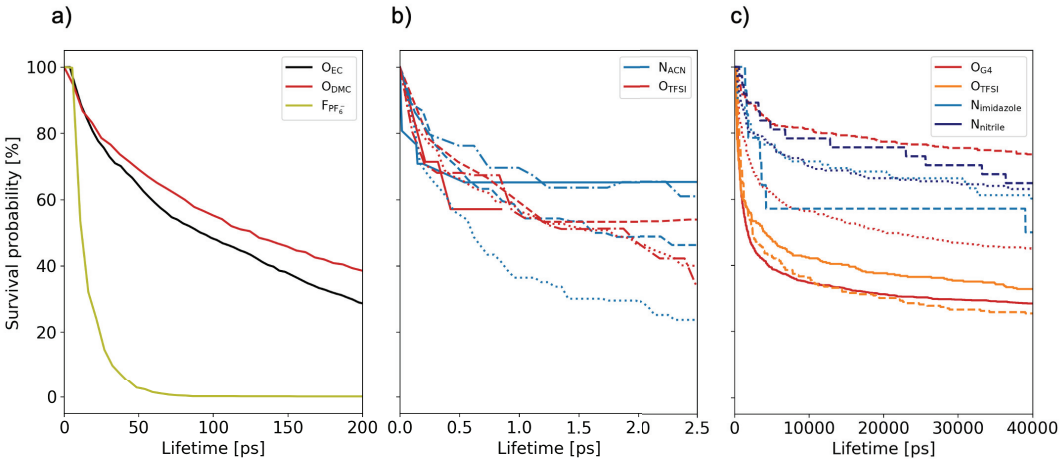


Figure 4.5: Survival curves for the residence times of Li-X coordination bonds, with X indicated in the legend: a) LP30, b) LiTFSI in ACN, low concentration (solid) to HCE (dotted), c) LiTFSI/LiTDI in G4 pure TFSI (solid) to pure TDI (dotted).

Considering each case in more detail, we see that in LP30, the solvents have much longer residence times than the anion, in contrast to what may be expected based on the overrepresentation of PF_6^- in the Li^+ 1st solvation shell (Fig. 4.4a). However, PF_6^- residence times are much longer if considered based on 4.4 Å distance between Li^+ and P, rather than based on the distance between Li and F, indicating the anions to rotate while remaining close to Li^+ [104].

A good solvent is by definition able to dissolve the salt, often due to high electric permittivity. Since in LP30 and other relatively dilute electrolytes the solvent molecules completely dominate the stoichiometry, there are many opportunities for them to "steal" away a Li^+ ion from a coordinated anion. On the other hand, anions are quite rare and therefore seldom attacks any given solvation shell, enabling relatively long-lived coordination bonds between cations and solvent molecules. Of course, solvents can also be exchanged in the solvation shell, but the thermodynamic driving force for such exchanges is likely much weaker.

This reasoning also holds up for the residence times observed in LiTFSI in ACN (Fig. 4.5b). At the lowest concentration, the solvent has longer residence time than the anion, in spite of the huge solvation shell overrepresentation of the latter (Fig. 4.4b). For the most concentrated electrolyte, the order is

reversed due to the combined effect of an ample supply of anions to disrupt solvent coordination, and the lack of a continuous solvent phase to break up ion pairs. That TFSI is more strongly bound to Li^+ than the solvents can also be seen in that it has roughly equal residence time across the concentration range with the exception of the very lowest concentration, where entropic effects simply dominate.

The stability of ACN exhibits a complex behaviour, which we have discussed in greater depth in paper IV. The ACN residence time decreases slightly from the lowest to the second lowest concentration due to greater structural diversity, which has its maximum around the second lowest concentration due to a higher degree of ion interaction enabling more numerous distinct local structures and transitions between them, which grows quadratically with the diversity. The second most concentrated electrolyte has more or less the same stability as the least concentrated one. At this point the diversity is less and there is still a continuous solvent phase. Finally, from the second highest to the highest concentration, the residence time decreases dramatically, since isolated ACN molecules are quite ineffective at breaking up ion-ion bonds. The residence times we obtain are *ca.* 2 orders of magnitude different than published results from classical MD [41]. Note that this difference is not as dramatic as it sounds due to the exponential dependence on energy barriers discussed above, as well as methodological differences. We believe that the concentration ranges are valid, though the absolute values should be interpreted with some caution.

We finally turn to LiTFSI and LiTDI in G4 (Fig. 4.5c). Here again the smaller changes of anion stability with the composition indicates that cation-anion interaction is in general stronger than cation-solvent interaction, as also clear from the solvation shell overrepresentation (Fig. 4.4). It is also clear from Fig. 4.5 that the Li-TDI interaction is stronger in the condensed phase than Li-TFSI interaction. The nitrile nitrogens bind more stably to cations than do the imidazole nitrogens in the mixed electrolyte, while their stabilities seem to be identical, and near the average of the two for the pure LiTDI system, likely a consequence of both kinds of bonds taking part in forming the percolating network.

The most interesting, and hardest to explain, residence time trend is again that of the solvent. The solvent residence time goes from relatively short for the pure LiTFSI electrolyte, to extremely long for the mixed electrolyte, and intermediate for the pure LiTDI electrolyte. In the pure LiTFSI electrolyte, the Li-G4 residence time is slightly shorter than that of Li-TFSI, indicating the slightly greater interaction strength of the latter. At the other extreme, for the

pure LiTDI electrolyte, the ion disproportionation gives rise to domains where G4 mostly does not have to compete locally with TDI for Li^+ coordination, giving it greater stability. Nonetheless, if a solvated cation happens upon an anionic region the solvent molecule faces severe competition. The rigid cationic network may even function as a molecular sieve for separating Li^+ ions from their solvent. In the mixed electrolyte, the disproportionation still tends to locally increase solvation stability, but here there is no percolating network to form an effective sieve. At the same time, the lower concentration of TFSI ions with which the G4 chains have to compete compared to in the pure LiTFSI electrolyte also acts to increase the solvent residence time. Finally, an additional possibility is that the relative amounts of TFSI to solvent in the mixed system is such that the TFSI ions can be accommodated without competition, whereas that may not be the case in the pure LiTFSI system.

4.4 Diffusivity and Transport Mechanisms

We finally turn to the transport properties and mechanisms of ions and to some extent also solvent molecules. We generally find that solvents have higher diffusivity than ions. Furthermore, vehicular cation transport is facilitated by solvation and hindered by anion coordination due to the greater friction experienced by the electrically charged anion compared to the solvent molecules. Li^+ tends to have a larger fraction of non-vehicular transport than either solvents or anions due to its ever-present solvation shell. The difference is naturally less pronounced for species that interact to a great extent or with rapid dynamics with Li^+. Both anions and especially the normally faster-moving solvent molecules tend to be slowed down by coordination to a Li^+ ion at low concentrations, since this makes them part of a larger and more slow-moving vehicle than in their uncoordinated form.

Comparing the species diffusivities in the different electrolytes (Fig. 4.6), we again find widely different orders of magnitude, in the opposite ordering to the residence times, indicating a connection between the rate of structural dynamics and species diffusivities.

Considering the individual electrolytes, in LP30, we observe Li^+ and PF_6^- to have almost the same diffusivity, giving $t^+ \approx 0.5$, which is greater than observed in experimental studies, *ca.* 0.2-0.3, but seems to be a common feature of computational studies for unexplained reasons (Fig. 4.6a). Both solvents have roughly the double diffusivity.

With respect to non-vehicular Li^+ transport, addition and removal of a ligand give similar contributions to the diffusivity. Exchanges of solvent

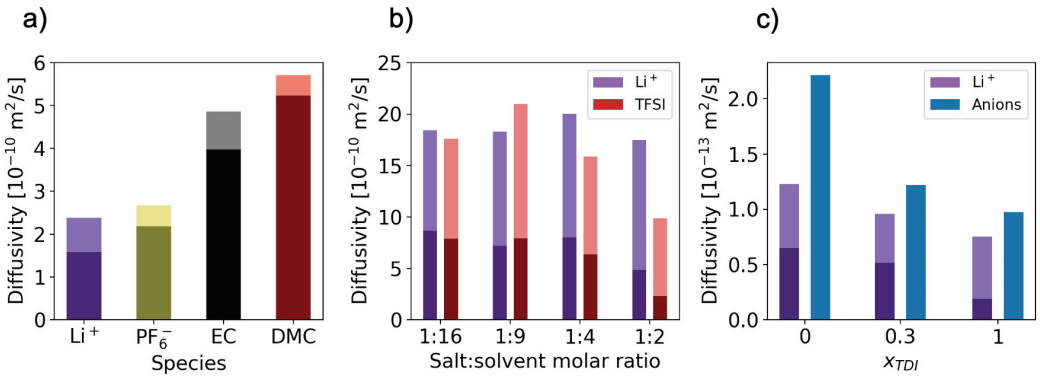


Figure 4.6: Diffusivities of Li^+, anions and solvent species, also decomposed into vehicular and non-vehicular parts: a) LP30, b) LiTFSI in ACN, c) LiTFSI/LiTDI in G4.

molecules give a much greater push to the Li^+ ion on average than anion exchange, yielding a greater contribution. EC exchange has the biggest contribution due to its higher frequency than DMC exchange.

In LiTFSI in ACN, increasing salt concentration leads to increasing t^+ and decreasing vehicular fraction of transport for both Li^+ and TFSI (Fig. 4.6b). Again, solvent diffusivity is roughly double that of the cation, and we observe $t^+ \approx 0.5$ at lower concentrations, at odds with experimental results and common knowledge. Interestingly, the only clear change in diffusivity we see over the observed concentration range is a marked drop in TFSI diffusivity from the second highest to the highest concentration.

While the transport mechanisms at lower concentrations is more akin to those observed for LP30, a detailed investigation of the mechanism in the most concentrated electrolyte (Paper I) showed that while both cations and anions participate almost fully in forming the percolating network (Fig. 4.1c), the TFSI anions are more immobilised since the tend to be fixed at several different coordination sites. The flexibility of TFSI facilitates structural reorganisation enabling solvated Li^+ ions to move in the network with hardly any concomitant movement of the TFSI centre-of-mass.

Turning finally to the G4 based electrolytes LiTFSI/LiTDI in G4: TFSI forms moderately sized flexible chains of solvated cations alternating with anions. TDI forms a rigid anionic percolating network and smaller, cationic structures, often devoid of direct anion interaction. Very low diffusivities, with clear transition from ballistic to structurally constrained diffusion. TDI

addition seems to lower the diffusivity of all species monotonously, but the disproportionation of cations and anions that it causes leads to greater average net charge of small and mobile connected components, thus giving a boost to conductivity and arguably cation transport number, albeit not much of a boost to cation transport number.

Chapter 5

Conclusions and Future Work

Using our newly developed framework for DSD across different classes of HCEs and more conventional electrolytes, we have found that ion transport in liquid battery electrolytes is generally carried by many different species, which additionally often change on the ps-ns scale, *i.e.* on the same range relevant for the transport processes themselves. This applies not only to HCEs, but also to traditional LIB electrolytes. We have also found that HCEs are not necessarily devoid of free solvent, but that their distinctive behaviour rather arises from the lack of free ions, and of a continuous solvent phase, as also argued in a recent review [34]. We have furthermore successfully used this framework to elucidate detailed transport mechanisms, including ones mediated by rearrangements of percolating disordered networks and ion disproportionation, as well as more conventional mechansims such as vehicular transport. Based on these insights, a full understanding of the transport mechanisms in battery electrolytes in general requires detailed knowledge of both the structure and dynamics of the varied species.

Traditional transport models such as the Stokes-Einstein or Maxwell-Stefan equations are therefore insufficient in two regards; first, they fail to account explicitly for speciation, treating instead effective properties of the "naked" ions and solvent molecules, and second, they implicitly assume the transported species to be infinitely long-lived. While such approaches can often be predictively valid with empirically fitted parameters, they cannot describe the actual molecular-scale dynamics. However, based on the detailed knowledge of speciation and dynamics gained from applying our novel analysis methods, these approaches can still be very useful, but based on the actual structural species rather than an ideal case assumed *a priori*.

We found that the Stokes-Einstein equation applied to the dynamically discovered structures is a decent first approximation of the vehicular part of Li^+ transport, but that higher order models such as Maxwell-Stefan need to be introduced to account for the different frictions felt by the different species. Instead of the $n(n-1)/2$ independent friction coefficients naively required

in a Maxwell-Stefan equation for n species, it should be possible to develop models to compute these from the shape and composition of the surface of each structural species.

Along these and similar lines, we envision DSD and our implementation of it, CHAMPION, to be useful as one of the bottom layers in a MSM framework. Due to its both qualitative and quantitative nature, it can be useful both in informing the development of more macroscopic models of structure and dynamical processes, and in providing these with the needed parameters without needing to fit macroscopic measurements which add scant understanding. Databases with output from many DSD analyses can be used to generate and test hypotheses for more general transport models, possibly aided by data mining.

DSD results could alternatively be fed into *e.g.* kinetic Monte Carlo approaches, or directly into *e.g.* Newman style modelling, depending on the complexity of the application. A MSM stack based on AIMD with non-empirical functionals and DSD is entirely a bottom-up approach, thus potentially eliminating the need for empirical parameters altogether.

Let us finally outline some short-term research directions we would like to explore. We would like to return to to the original vision of machine learning force fields to enable classical MD simulations with AIMD accuracy. This would form a layer between AIMD and DSD in the envisioned MSM stack. Whether we succeed in our approach to this problem, multiple groups have already launched successful ML frameworks for force field generation, including Schnet [105, 106], Gaussian Approximation Potentials [107, 108] and Behler-Parinnello neural networks [109, 110]. These tend to have computational cost intermediate between AIMD and classical MD, but they still enable much larger-scale simulations to be run compared to AIMD.

To date we have only applied our framework to bulk electrolytes. While understanding of the bulk is definitely highly valuable, the operation of electrochemical systems is also highly dependent on surface phenomena. Our methods could at least in principle just as well be applied to simulation trajectories near *e.g.* electrode interfaces. It would be very interesting to see how for instance an anion-blocking but Li^+ accepting strongly charged surface would affect the local speciation.

Furthermore, the framework should be applied to many more systems than to date. Obvious examples include electrolytes for Li-S, sodium-ion batteries and other novel battery chemistries, as well as related electrolyte concepts such as LHCEs, aqueous HCEs, ionic liquid-based electrolytes. It could also be applied to polymer-based electrolytes such as GPEs and SPEs, and elec-

trode slurries. Indeed, there is no reason to limit future applications of the framework to the battery field. it could in principle be used to increase the understanding of any material that is complex in terms of either structure or dynamics, a broad class that includes among other materials all soft matter, polymeric systems and glasses.

There are no clear limits to the number of different materials to which the methods presented here can be applied. Especially in combination with other tools, the sky is truly the limit. It is the author's sincerest hope that my work will be widely used and contribute to solving real-life problems, thus making the world a slightly better place.

Acknowledgements

A PhD thesis is quite an adventure, in both the exhilarating and excruciating senses of the word. While it has been mostly fun, I could never have made it through without an amazing cast of fellow travellers.

First of all, thank you Patrik, my main supervisor, for your uncanny ability to salvage clear and concise writing from a bloated pile of nonsense. Thank you for giving me this opportunity, for letting me follow my own visions, and for restoring my belief in myself when the going got tough. Thank you Alejandro, my assistant supervisor, for helping me explore the dizzying realms of multi-scale modelling. Thank you for always reminding me that there can never be too many pictures in a presentation. Thanks to my external collaborators: Piotr Jankowski, Oleg Borodin, Wujie Wang and Rafael Gomez-Bombarelli. Thanks to everyone at LRCS for hosting me for two weeks in 2016, and to the MIT groups we visited in 2019. Special thanks to Prof. Yang Shao-Horn for invitating us. And thanks to my funders for making it all possible: Energimyndigheten, H2020 HELIS, and Chalmers Area of Advance Materials Science (acknowledged in greater detail on the inside cover).

Thanks to all the great colleagues at KMF (yes, I'm keeping the 'K', sue me!), thank you for all the fun lunches and fikas, all unforgettable after-works, kräftskivor, julbord and conferences. There has never been a research division I would rather have been part of.

Thanks to all the inspiring teachers I have had along the way: Nial Olsson, for showing (and telling) the virtues of the Socratic method and of thinking for oneself. Thanks to Olle Olsson for sparking my love for physics, and for showing (not telling) that geeks can rock and rockers can count. Thanks to Igor Zoric for bringing the atoms to life like no one else with just a few kindergarten quantum mechanics tricks. And thank you Christophe Demaziére for helping me find my true calling in computational physics.

Thanks to Fabian, Bo, Johannes and Emil at Compular, and thanks to our friends at Chalmers Ventures and Chalmers Innovation. Let's make it happen!

And last, but very certainly not least: thanks from the bottom of my heart, to my wife Emily and son Elliot for putting up with my crazy work hours and supporting me throughout this endeavor. I love you! Thank you for your patience and your love, and for being here.

References

(1) Ritchie, H.; Roser, M. CO2 and Other Greenhouse Gas Emissions., `https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions` (accessed 11/19/2020).

(2) Nishi, Y. *The Chemical Record*, *1*, 406–413.

(3) Schott, B.; Püttner, A.; Müller, M. In *Advances in Battery Technologies for Electric Vehicles*, Scrosati, B., Garche, J., Tillmetz, W., Eds.; Woodhead Publishing Series in Energy; Woodhead Publishing: 2015, pp 35–54.

(4) The Nobel Prize in Chemistry 2019., `https://www.nobelprize.org/prizes/chemistry/2019/summary/`.

(5) Xu, K. *Chem. Rev.* **2004**, *104*, 4303–4418.

(6) Petibon, R.; Xia, J.; Ma, L.; Bauer, M. K. G.; Nelson, K. J.; Dahn, J. R. *J. Electrochem. Soc.* **2016**, *163*, A2571–A2578.

(7) Berg, H., *Batteries for Electric Vehicles: Materials and Electrochemistry*; Cambridge university press: 2015.

(8) Wu, Y., *Lithium-Ion Batteries: Fundamentals and Applications*; CRC Press: 2015; Vol. 4.

(9) Nitta, N.; Wu, F.; Lee, J. T.; Yushin, G. *Materials Today* **2015**, *18*, 252–264.

(10) Ji, X.; Nazar, L. F. *J. Mater. Chem.* **2010**, *20*, 9821–9826.

(11) Nazar, L. F.; Cuisinier, M.; Pang, Q. *MRS Bulletin* **2014**, *39*, 436–442.

(12) Choi, J. W.; Aurbach, D. *Nature Reviews Materials* **2016**, *1*, 16013.

(13) Greenwood, N. N.; Earnshaw, A., *Chemistry of the Elements*; Elsevier: 2012.

(14) Scheers, J.; Fantini, S.; Johansson, P. *Journal of Power Sources* **2014**, *255*, 204–218.

(15) Flamme, B.; Garcia, G. R.; Weil, M.; Haddad, M.; Phansavath, P.; Ratovelomanana-Vidal, V.; Chagnes, A. *Green Chemistry* **2017**, *19*, 1828–1849.

(16) Suo, L.; Borodin, O.; Gao, T.; Olguin, M.; Ho, J.; Fan, X.; Luo, C.; Wang, C.; Xu, K. *Science* **2015**, *350*, 938–943.

(17) Etacheri, V.; Marom, R.; Elazari, R.; Salitra, G.; Aurbach, D. *Energy & Environmental Science* **2011**, *4*, 3243–3262.

(18) Peled, E. *J. Electrochem. Soc.* **1979**, *126*, 2047–2051.

(19) Peled, E.; Menkin, S. *J. Electrochem. Soc.* **2017**, *164*, A1703–A1719.

(20) Guyomard, D.; Tarascon, J. M. *J. Electrochem. Soc.* **1993**, *140*, 3071.

(21) Tarascon, J. M.; Guyomard, D. *Solid State Ionics* **1994**, *69*, 293–305.

(22) Guyomard, D.; Tarascon, J. M. *J. Electrochem. Soc.* **1992**, *139*, 937.

(23) Dahn, J. R.; von Sacken, U.; Juzkow, M. W.; AlJanaby, H. *J. Electrochem. Soc.* **1991**, *138*, 2207.

(24) Fong, R.; von Sacken, U.; Dahn, J. R. *J. Electrochem. Soc.* **1990**, *137*, 2009.

(25) Zhang, S. S. *Journal of Power Sources* **2006**, *162*, 1379–1394.

(26) Meazah Haregewoin, A.; Sorsa Wotango, A.; Hwang, B.-J. *Energy & Environmental Science* **2016**, *9*, 1955–1988.

(27) Abe, K.; Yoshitake, H.; Kitakura, T.; Hattori, T.; Wang, H.; Yoshio, M. *Electrochimica Acta* **2004**, *49*, 4613–4622.

(28) Hu, Y.; Kong, W.; Li, H.; Huang, X.; Chen, L. *Electrochemistry Communications* **2004**, *6*, 126–131.

(29) Liu, J.; Chen, Z.; Busking, S.; Amine, K. *Electrochemistry Communications* **2007**, *9*, 475–479.

(30) Jankowski, P.; Wieczorek, W.; Johansson, P. *J Mol Model* **2017**, *23*, 6.

(31) McKinnon, W. R.; Dahn, J. R. *J. Electrochem. Soc.* **1985**, *132*, 364–366.

(32) Yamada, Y.; Yamada, A. *Journal of The Electrochemical Society* **2015**, *162*, A2406–A2423.

(33) Yamada, Y.; Wang, J.; Ko, S.; Watanabe, E.; Yamada, A. *Nat Energy* **2019**, *4*, 269–280.

(34) Borodin, O.; Self, J.; Persson, K. A.; Wang, C.; Xu, K. *Joule* **2020**, *4*, 69–100.

(35) Yamada, Y.; Furukawa, K.; Sodeyama, K.; Kikuchi, K.; Yaegashi, M.; Tateyama, Y.; Yamada, A. *Journal of the American Chemical Society* **2014**, *136*, 5039–5046.

(36) Jeong, S.-K.; Inaba, M.; Iriyama, Y.; Abe, T.; Ogumi, Z. *Electrochemical and Solid-State Letters* **2003**, *6*, A13.

(37) Jeong, S.-K.; Inaba, M.; Iriyama, Y.; Abe, T.; Ogumi, Z. *Journal of Power Sources* **2008**, *175*, 540–546.

(38) Matsumoto, K.; Inoue, K.; Nakahara, K.; Yuge, R.; Noguchi, T.; Utsugi, K. *Journal of Power Sources* **2013**, *231*, 234–238.

(39) Yoshida, K.; Nakamura, M.; Kazue, Y.; Tachikawa, N.; Tsuzuki, S.; Seki, S.; Dokko, K.; Watanabe, M. *J. Am. Chem. Soc.* **2011**, *133*, 13121–13129.

(40) Seo, D. M.; Borodin, O.; Han, S.-D.; Ly, Q.; Boyle, P. D.; Henderson, W. A. *Journal of The Electrochemical Society* **2012**, *159*, A553–A565.

(41) Seo, D. M.; Borodin, O.; Han, S.-D.; Boyle, P. D.; Henderson, W. A. *Journal of The Electrochemical Society* **2012**, *159*, A1489–A1500.

(42) Seo, D. M.; Borodin, O.; Balogh, D.; O'Connell, M.; Ly, Q.; Han, S.-D.; Passerini, S.; Henderson, W. A. *Journal of The Electrochemical Society* **2013**, *160*, A1061–A1070.

(43) Han, S.-D.; Borodin, O.; Allen, J. L.; Seo, D. M.; McOwen, D. W.; Yun, S.-H.; Henderson, W. A. *Journal of The Electrochemical Society* **2013**, *160*, A2100–A2110.

(44) Han, S.-D.; Borodin, O.; Seo, D. M.; Zhou, Z.-B.; Henderson, W. A. *Journal of The Electrochemical Society* **2014**, *161*, A2042–A2053.

(45) Borodin, O.; Han, S.-D.; Daubert, J. S.; Seo, D. M.; Yun, S.-H.; Henderson, W. A. *Journal of The Electrochemical Society* **2015**, *162*, A501–A510.

(46) McOwen, D. W.; Seo, D. M.; Borodin, O.; Vatamanu, J.; Boyle, P. D.; Henderson, W. A. *Energy & Environmental Science* **2014**, *7*, 416–426.

(47) Suo, L.; Hu, Y.-S.; Li, H.; Armand, M.; Chen, L. *Nat Commun* **2013**, *4*, 10.1038/ncomms2513, 1481.

(48) Forsyth, M.; Yoon, H.; Chen, F.; Zhu, H.; MacFarlane, D. R.; Armand, M.; Howlett, P. C. *J. Phys. Chem. C* **2016**, *120*, 4276–4286.

(49) Flores, E.; Åvall, G.; Jeschke, S.; Johansson, P. *Electrochimica Acta* **2017**, *233*, 134–141.

(50) Gustav Åvall, Patrik Johansson *Journal of Chemical Physics* **Accepted May 8th, 2020**, *152*.

(51) Kühnel, R.-S.; Reber, D.; Battaglia, C. *ACS Energy Lett.* **2017**, *2*, 2005–2006.

(52) Okoshi, M.; Chou, C.-P.; Nakai, H. *J. Phys. Chem. B* **2018**, *122*, 2600–2609.

(53) Lundgren, H.; Scheers, J.; Behm, M.; Lindbergh, G. *Journal of The Electrochemical Society* **2015**, *162*, A1334–A1340.

(54) Nilsson, V.; Younesi, R.; Brandell, D.; Edström, K.; Johansson, P. *Journal of Power Sources* **2018**, *384*, 334–341.

(55) Nilsson, V.; Kotronia, A.; Lacey, M.; Edström, K.; Johansson, P. *ACS Appl. Energy Mater.* **2020**, *3*, 200–207.

(56) Nilsson, V.; Bernin, D.; Brandell, D.; Edström, K.; Johansson, P. *ChemPhysChem* **2020**, *n/a*, DOI: 10.1002/cphc.202000153.

(57) Kameda, Y.; Saito, S.; Saji, A.; Amo, Y.; Usuki, T.; Watanabe, H.; Arai, N.; Umebayashi, Y.; Fujii, K.; Ueno, K.; Ikeda, K.; Otomo, T. *J. Phys. Chem. B* **2020**, DOI: 10.1021/acs.jpcb.0c08021.

(58) Dokko, K.; Tachikawa, N.; Yamauchi, K.; Tsuchiya, M.; Yamazaki, A.; Takashima, E.; Park, J.-W.; Ueno, K.; Seki, S.; Serizawa, N.; Watanabe, M. *J. Electrochem. Soc.* **2013**, *160*, A1304.

(59) Ueno, K.; Park, J.-W.; Yamazaki, A.; Mandai, T.; Tachikawa, N.; Dokko, K.; Watanabe, M. *J. Phys. Chem. C* **2013**, *117*, 20509–20516.

(60) Suo, L.; Borodin, O.; Sun, W.; Fan, X.; Yang, C.; Wang, F.; Gao, T.; Ma, Z.; Schroeder, M.; von Cresce, A.; Russell, S. M.; Armand, M.; Angell, A.; Xu, K.; Wang, C. *Angewandte Chemie* **2016**, *128*, 7252–7257.

(61) Yang, C.; Chen, J.; Qing, T.; Fan, X.; Sun, W.; von Cresce, A.; Ding, M. S.; Borodin, O.; Vatamanu, J.; Schroeder, M. A.; Eidson, N.; Wang, C.; Xu, K. *Joule* **2017**, *1*, 122–132.

(62) Zhu, J.; Xu, Y.; Fu, Y.; Xiao, D.; Li, Y.; Liu, L.; Wang, Y.; Zhang, Q.; Li, J.; Yan, X. *Small* **2020**, *16*, 1905838.

(63) Baskin, A.; Prendergast, D. *J. Phys. Chem. C* **2016**, *120*, 3583–3594.

(64) Chen, S.; Zheng, J.; Mei, D.; Han, K. S.; Engelhard, M. H.; Zhao, W.; Xu, W.; Liu, J.; Zhang, J.-G. *Advanced Materials* **2018**, *30*, 1706102.

(65) Drvaric Talian, S.; Jeschke, S.; Vizintin, A.; Pirnat, K.; Aron, I.; Aquilanti, G.; Johansson, P.; Dominko, R. *Chem. Mater.* **2017**, DOI: 10.1021/acs.chemmater.7b03654.

(66) Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids: Second Edition*; Oxford University Press: 2017; 641 pp.

(67) Newman, J. S.; E, T.-A. K., *Electrochemical Systems*; John Wiley & Sons: 2004.

(68) Doyle, M.; Newman, J. *Electrochimica Acta* **1995**, *40*, 2191–2196.

(69) Lai, W.; Ciucci, F. *Electrochimica Acta* **2011**, *56*, 4369–4377.

(70) Krishna, R.; Wesselingh, J. A. *Chemical Engineering Science* **1997**, *52*, 861–911.

(71) Nyman, A.; Behm, M.; Lindbergh, G. *Electrochimica Acta* **2008**, *53*, 6356–6365.

(72) Krachkovskiy, S. A.; Bazak, J. D.; Fraser, S.; Halalay, I. C.; Goward, G. R. *J. Electrochem. Soc.* **2017**, *164*, A912–A916.

(73) Krachkovskiy, S.; Dontigny, M.; Rochon, S.; Kim, C.; Trudeau, M. L.; Zaghib, K. *J. Phys. Chem. C* **2020**, *124*, 24624–24630.

(74) Borodin, O.; Olguin, M.; Ganesh, P.; Kent, P. R. C.; Allen, J. L.; Henderson, W. A. *Phys. Chem. Chem. Phys.* **2015**, *18*, 164–175.

(75) Marx, D.; Hutter, J., *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: 2009.

(76) Kaminski, G.; Jorgensen, W. L. *J. Phys. Chem.* **1996**, *100*, 18010–18013.

(77) Ponder, J. W.; Case, D. A. In *Advances in Protein Chemistry*; Protein Simulations, Vol. 66; Academic Press: 2003.

(78) Bedrov, D.; Borodin, O.; Li, Z.; Smith, G. D. *J. Phys. Chem. B* **2010**, *114*, 4984–4997.

(79) Latz, A.; Zausch, J.; Iliev, O. In *Numerical Methods and Applications*, ed. by Dimov, I.; Dimova, S.; Kolkovska, N., Springer Berlin Heidelberg: 2011, pp 329–337.

(80) Schammer, M.; Horstmann, B.; Latz, A. Theory of Transport in Highly Concentrated Electrolytes., Comment: 16 pages, 9 figures, http://arxiv.org/abs/2010.14915 (accessed 11/18/2020).

(81) Quiroga, M. A.; Xue, K.-H.; Nguyen, T.-K.; Tulodziecki, M.; Huang, H.; Franco, A. A. *Journal of The Electrochemical Society* **2014**, *161*, E3302–E3310.

(82) Storey, B. D.; Bazant, M. Z. *Phys. Rev. E* **2012**, *86*, 056303.

(83) McEldrew, M.; Goodwin, Z. A. H.; Bi, S.; Bazant, M. Z.; Kornyshev, A. A. *J. Chem. Phys.* **2020**, *152*, 234506.

(84) Gavish, N.; Elad, D.; Yochelis, A. *J. Phys. Chem. Lett.* **2018**, *9*, 36–42.

(85) Thijssen, J. M., *Computational Physics*; Cambridge University Press: 2007.

(86) Nolting, W., *Fundamentals of Many-Body Physics: Principles and Methods*; Springer-Verlag: Berlin Heidelberg, 2009.

(87) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(88) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(89) Perdew, J. P.; Schmidt, K. *AIP Conference Proceedings* **2001**, *577*, 1–20.

(90) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(91) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

(92) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *Journal of Computational Chemistry*, *7*, 230–252.

(93) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, doi: 10.1021/ja9621760, 11225–11236.

(94) Einstein, A. *Annalen der Physik* **1905**, *vol. 4, t. 17*.

(95) Illustration by Wikimedia user Pasafr under CC-BY-SA 4.0 (`https://creativecommons.org/licenses/by-sa/4.0/deed.en`): Comparison between Gradient Method (White) and Conjugate Gradient Method (Red)., 2016.

(96) Kramer, O., *Genetic Algorithm Essentials*; Studies in Computational Intelligence; Springer International Publishing: 2017.

(97) Rasmus Andersson; Fabian Årén; Patrik Johansson Dynamic Bond Detection. pat., application 2051245-5 (SE).

(98) Johansson, P.; Gejji, S. P.; Tegenfeldt, J.; Lindgren, J. *Electrochimica Acta* **1998**, *43*, 1375–1379.

(99) Dranka, M.; Niedzicki, L.; Kasprzyk, M.; Marcinek, M.; Wieczorek, W.; Zachara, J. *Polyhedron* **2013**, *51*, 111–116.

(100) Jankowski, P.; Dranka, M.; Wieczorek, W.; Johansson, P. *J. Phys. Chem. Lett.* **2017**, *8*, 3678–3682.

(101) Yamada, Y.; Usui, K.; Chiang, C. H.; Kikuchi, K.; Furukawa, K.; Yamada, A. *ACS Applied Materials & Interfaces* **2014**, *6*, 10892–10899.

(102) Xu, K. *Chem. Rev.* **2014**, *114*, 11503–11618.

(103) Von Wald Cresce, A.; Borodin, O.; Xu, K. *J. Phys. Chem. C* **2012**, *116*, 26111–26117.

(104) Chen, J. et al. *Nature Energy* **2020**, *5*, 386–397.

(105) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. *J. Chem. Phys.* **2018**, *148*, 241722.

(106) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. *Chem. Mater.* **2019**, *31*, 3564–3572.

(107) Bartók, A. P. Gaussian Approximation Potential: An Interatomic Potential Derived from First Principles Quantum Mechanics., Comment: PhD thesis, University of Cambridge 2009, `http://arxiv.org/abs/1003.2817` (accessed 01/31/2019).

(108) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction., Comment: 3 figures, 20 pages, `http://arxiv.org/abs/1502.01366` (accessed 01/31/2019).

(109) Behler, J.; Parrinello, M. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(110) Smith, J. S.; Isayev, O.; Roitberg, A. E. *Chem. Sci.* **2017**, *8*, 3192–3203.